CRANFIELD UNIVERSITY

MARTIN DOUGLAS

The Social Reality of Initiatives which Pursue Insight from Data

School of Management

PhD
Academic Year: 2010 - 2016

Supervisors: Prof Joe Peppard/ Dr Stan Maklan
March 2016

ABSTRACT

While (big) data promises immense opportunity, initiatives focused on using data to pursue insight have mixed outcomes. The Management Support Systems (MSS) model summarises what we currently understand within Information Systems (IS) about the implementation and use of systems to improve organisations' use of data. Adopting an ethnographic approach to observe how practitioners in two contrasting organisations actually generate insight from data, this research challenges the implicit information processing and implementation logics of the MMS model. The pragmatic messiness of pursuing insight is described in two monographs, which reveal the socially constructed nature of data in relation to phenomena, and the importance of data engagement to produce insight. Given that this PhD study also seeks to generate insight from data, it is compared and contrasted reflexively to the two cases observed. While the inquiry logic pursued in this study was made explicit, and was regularly reviewed and challenged, the two cases left this largely implicit. The use of tools is shown to facilitate and constrain inquiry, with related data acting as boundary objects between the different practitioner groups involved. An explanatory framework is presented and used to suggest various enhancements to the MSS model. First, the Problem Space is reframed to reflect the distinct, though interdependent logics involved in inquiry versus realising envisaged benefits from insights. Second, the MSS artefact itself is contextualised and Data Engagement rather than MSS or Tool Use is positioned as central. Third, Data are disentangled from the wider MSS artefact, as a critical, distinct construct. Fourth, an Alignment construct is introduced to address the boundary spanning nature of data initiatives. The thesis also highlights the value of using Wenger's (1998) Communities of Practice (CoP) situated learning framework to study data initiatives, and the related value of mapping groups as a technique for further development. Some questions are provided for practitioners to gain a better understanding of data initiatives. Wider implications are also noted for the socio-material theorising of Data, and distinguishing between Data, Information and Knowledge concepts within the IS discipline.

Keywords:

Data, Insight, Knowledge, Knowing, Situated Learning, Communities of Practice, Sociomateriality

ACKNOWLEDGEMENTS

While recognising the importance of *owning* my thesis, this section seeks to recognise and thank various people and organisations who have helped me to realise this research effort over several years.

First, I'd like to thank Joe, my supervisor, who always knew there was a PhD in there somewhere… even when its size and shape wasn't all that clear at the outset of the adventure. His continuing to supervise me, on a *pro bono* basis after leaving Cranfield University for the ESMT in Berlin, is much appreciated. His patient encouragement throughout, and knack for identifying what key things to focus on doing or improving next, were invariably helpful. I'd also like to thank Stan for gamely stepping in as co-supervisor, when this isn't his field, and to Deirdre for her ongoing encouragement as panel chair. Her exhortation to use the first person in connection with my work rather than my default collective language will stay with me.

The start of this journey must also be acknowledged. The University of Cape Town's (UCT) Accounting Department and its IS section, initially fostered my interest in the IS field many years ago now. Special mention here goes to Keith Mattison, who encouraged me as tutor in the field early on, as well as Peter Lay and Derek Smith. I'd also like to thank them for very kindly hosting me for three months in 2014 as a visiting student. Special mention here goes to Prof. Tessa Minter, for facilitating my visit, as well as for her unstinting encouragement, understanding and friendship. Another special mention to Daphne Barends for so kindly letting me stay with her during my UCT visits.

KPMG LLP also kindly provided bursary funding for the first few years of the PhD and allowed me a flexible working arrangement, which made a big difference given the time commitment required for such research. Julian Schwarzenberg, BCS Data Management SIG chair, provided me continued encouragement about the importance of the human aspects of Data initiatives, and introduced me to one of my cases.

Special thanks go to the participants in both case settings, especially those I worked particularly closely with in each setting. You were instrumental to making this a success and making it fun and enjoyable too – you provided valuable, open, nuanced insights into what was actually going on, especially at an interpersonal level, without which such research would not be nearly as useful or interesting.

Finally, thanks to all my friends and family for being so understanding and supportive over such a long period. The support has even extended to providing periodic food parcels and vitamin supplements. I have disappeared from view for many of you for long periods of time and look forward to catching up about something other than the state of my PhD in future. Thanks too to several new friends made from my PhD cohort – a special group who have sustained each other over several years with humour, fun and really constructive review comments and suggestions. This includes Donna Ladkin, our research methods course chair, who got us off to a particularly memorable start.

Dissemination of this research to date

This research has been disseminated and presented in part as follows:

Academic Conference Papers

Douglas, M.D. and Peppard, J. (2013) Theorizing Data, information and Knowledge constructs and their inter-relationship, in *Proceedings of 18th UKAIS Conference on Information Systems (UKAIS: 2013)*, Oxford, United Kingdom, available at http://aisel.aisnet.org/ukais2013/

Douglas, M.D. and Peppard, J. (2014) Adopting a situated learning framework for (big) data projects, in *Proceedings of International Data and Information Management Conference (IDIMC 2014)*, Loughborough, United Kingdom, available at http://idimc.org/2014-conference-3/conference-proceedings/
(Awarded best conference paper)

Practioner seminar

Adopting a situated learning framework for (big) data projects
British Computer Society – Data Special Interest Group Meeting – October 2014

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

BCS	British Computer Society
BI	Business Intelligence
BIM	Building Information Management
CAD	Computer Aided Design
CDO	Chief Data Officer
CEO	Chief Executive Officer
CIO	Chief Information Officer
COBie	Construction Operations Building information exchange
CoP	Communities of Practice
CRM	Customer Relationship Management
DAMA	Data Management Association
DW	Data Warehouse
DSS	Decision-Support Systems
EIS	Executive Information Systems
GIS	Geographical Information System
ICT	Information Communications Technology
IM	Information Management
IS	Information Systems
ISACA	Information Systems Audit and Control Association
IT	Information Technology
KM	Knowledge Management
KMS	Knowledge Management Systems
KPI	Key Performance Indicator
MSS	Management Support Systems
OL	Organisational Learning
PR	Public Relations
RAG	Red, Amber, Green
R&D	Research and Development
SMART	Specific, Measurable, Action-oriented, Realistic and Time-bound
SPSS	Software Package for Statistical Analysis
SQL	Structured Query Language
XBRL	eXtensible Business Reporting Language
XML	eXtensible Markup Language

1 INTRODUCTION

The rapidly growing availability of data, including new forms of data (e.g. social media and location data), is widely seen as a significant opportunity to derive new insights to create related value for organisations (Davenport: 2014), especially when exploited using new tools and analytical techniques (Chen, Chiang & Storey: 2012). Related organisational capabilities are widely regarded as an important dimension of corporate competitiveness (Kettinger & Marchand: 2011, Davenport: 2009). While this has prompted many data initiatives, implementing a variety of data analytics technologies (Chen et al.: 2012, Ranjan & Bhattnagar: 2011, Bose: 2009), many investments and data initiatives result in mixed outcomes (Marchand & Peppard: 2013, Yeoh & Koronios: 2010, Wixom & Watson: 2001, Cooper, Watson & Wixom: 2000), often generating a wealth of data but a poverty of insight.

This is explored in section 1.1 below which provides background and context to the problem my research seeks to address. Section 1.2 then briefly sets out how the research addresses the challenges outlined and how the thesis argument is structured.

1.1 Background and Business Problem

One of the most profound impacts of information technology (IT) is the continuing, rapid increase in the volume and nature of digital data being captured, stored and made available for use (Hemp: 2009, Davenport et al.: 2001). This has given rise to new terms such as ‘Big Data’ in recognition of this trend, with social media and location data as examples of new kinds of digital data being captured.

The rapidly growing availability of such data is widely argued to be a significant opportunity for organisations to derive new insights from such data and exploit them to create value. Indeed, many economists (Foss & Mahnke: 2003) argue that the value of organising (and therefore of organisations) is based on the principle of exploiting information or knowledge asymmetries (i.e. specialisation). This is reflected in thinking of organisations as knowledge

‘engines’ or ‘information processors’ operating in a knowledge economy. Several examples exist of how this abundance of data is being exploited successfully, especially (but not exclusively) by ‘pure-play’ Internet companies such as Google and Amazon (Chen et al.: 2012, Cooper et al.: 2000).

Perhaps on the basis of such companies’ success, information and knowledge related organisational capabilities are now widely regarded as being important to corporate success and important dimensions of corporate competitiveness (Davenport: 2009, Marchand, Kettinger & Rollins: 2001, Davenport, Harris & Morison: 2001). This puts pressure on managers to demonstrate an appropriate response, prompting continued related investment in new systems and technology adoption (Davenport et al: 2001).

Recently much corporate IT investment has focused on implementing Customer Relationship Management (CRM), Business Intelligence (BI), Data Warehouse (DW) and Data Analytics systems. Such investment seeks to exploit both organisational and external data, knowledge management and collaboration tools in order to generate new insight. Given the focus of the thesis on generating insight from data, and Data Analytics being a relatively new development, this is explored below to provide further context.

1.1.1 Data Analytics

The term *data analytics* has emerged in the business press and been popularised by academics such as Davenport et al. (2001). They argue there is a significant opportunity to adopt more analytical approaches to management and decision-making, grounded in the use of the growing abundance of objective data available.

Chen et al. (2012) identify an evolution of BI and data analytics based on differentiated characteristics of the kind of data used and related analytical techniques adopted. They identify three such phases, starting with structured data, progressing to unstructured data and content, and finally addressing sensor and mobile content. They also recognise the differentiated adoption of

data analytics within fields or domains rather than a uniform trend, and highlight the following applications as particularly promising:

- eCommerce and market intelligence
- eGovernment and politics
- Science and Technology
- Smart Health and Wellbeing
- Security and Public Safety

With a different, more commercial starting point and focus, the McKinsey Global Institute (Manyika et al.: 2011), argue that different industry sectors have different opportunities to capture value from big data and related analytics. They argue that the scale of value opportunity depends on the following factors: talent, IT intensity, data-driven mindset and data availability within each sector.

Turning to the academic literature for a description of data analytics, Bose (2009) describes *advanced* analytics as follows:

> *"Advanced analytics is a general term which simply means applying various advanced analytic techniques to data to answer questions or solve problems. It is not a technology in and of itself, but rather, a group of tools that are used in combination with one another to gain information, analyze that information, and predict outcomes of the problem solutions. Data integration and data mining are the basis for advanced analytics" (p.156)*

Bose (2009) highlights the growing collection, integration and use of structured and unstructured data, as well as a shift from historical analysis towards predictive modelling of outcomes, which is identified as one of the differentiators of more *advanced* approaches. Bose provides the following diagram by way of overview description of BI architecture incorporating advanced analytics:

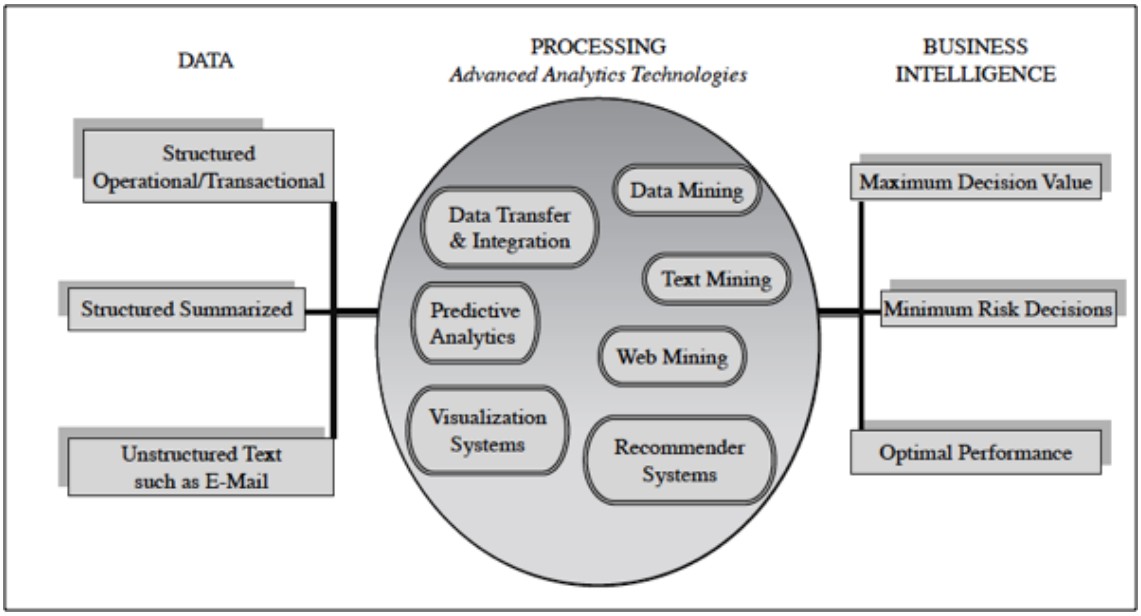


Figure 1-1 BI architecture incorporating Advanced Analytics (Bose: 2009: p.159)

This architecture in Figure 1-1 describes three broad components: collecting and integrating data, analytic technologies processing the data and systems to deliver the results to users, although collectively describing them as parts of a BI architecture. These processing technologies keep evolving and machine learning has more recently emerged and could be added to this set. While theirs is a useful characterisation, within the broader IS academic literature there seems to be a lack of clarity as to terms or consensus in describing BI systems (Arnott & Pervan: 2008, 2005).

Clark, Jones and Armstrong (2007) argue for the concept of Management Support Systems (MSS), as a broad class of similar and related systems which all support management decision-making and action, and can be studied together. While they don't specifically address Data Analytics, they include BI, Decision Support Systems (DSS), Executive Information Systems (EIS), Knowledge Management Systems (KMS) and their broad definition seems to encompass analytic systems and tools.

These systems are described broadly as supporting management decision makers by providing access to data, together with analytical capabilities with which to analyse such data, or assistance in analysing what they term a

problem space (Clark et al: 2007). They argue for considering and leveraging research related to these various areas in studying similar systems, presenting a common conceptual framework for such systems. This MSS model is examined in more detail in the Chapter 2 literature review.

1.1.2 Mixed outcomes and dominant logic

While considerable promise and opportunity has been identified, actual results from investments in such systems have been mixed across different implementation projects and organisations (Marchand & Peppard: 2013, Yeoh & Koronios: 2010, Jourdan, Rainer & Marshall: 2008, Wixom & Watson: 2001, Davenport et al: 2001, Cooper et al.: 2000). Concern about such mixed investment outcomes appears to be a special instance of a wider concern about IS implementation outcomes (Arvidsson, Holmström & Lyytinen: 2014, Thorp: 2003). Many efforts seek to improve the value or benefits derived from investments in tools and automated systems by enabling more effective organisational use of data, information and knowledge in terms of improved decision-making. However, managers and researchers seem to think about the use of data, information and knowledge quite straightforwardly, as illustrated by the following hierarchy about data and its relationship to knowledge (Rowley: 2007). This hierarchy is commonly refered to by practitioners and is also popular amongst IS academics (Kettinger & Li: 2010). The hierarchy below (see Figure 1-2) treats moving from data to information to knowledge as fairly straightforward and implicitly sequential.

Figure 1-2 The Data-Information-Knowledge-Wisdom pyramid (per Ackoff: 1989)

The current models for data analytics touched on in section 1.1.1, also seem to be driven or underpinned by a straightforward instrumentalist or processing logic: applying techniques or algorithms to data produces insights for decision-makers to make better decisions. This logic might be summarised as follows:

Data + Techniques = Insight

In terms of such logic, improved insight might then be achieved as follows:

More Data + **Better** Techniques = **More** Insight

To the extent that such improved techniques or algorithms are often automated or encoded within related software tools, this logic can also be expressed as follows:

More Data + **Better** Tools = **More** Insight

This is reflected in the emphasis on tool implementation encountered in the IS models and research focus outlined above. However, adopting such logic may unconsciously frame and bound thinking, ignoring important factors that contribute to the success of such initiatives, both to generate insight and to realise related benefits. In particular, the role of human agency and social context are hidden from view when adopting such logic or framing. For instance, managers and users of data tend not to be reflected in such logics. Also, the focus is on their improved decision-making rather than including wider learning benefits that may be achieved. The mixed outcomes from such initiatives may partly stem from shortcomings in such underpinning dominant logics.

1.1.3 The particular challenges faced by Exploratory Data projects

While one response to achieving better project outcomes has been to pay closer attention to benefits management (Ward & Daniel: 2012), this recommends identifying clear benefits desired, related dependencies, and ensuring appropriate accountabilities and attention are paid to achieving these. While such approaches hold considerable promise for improving projects where the desired outcomes and approaches to realise them are relatively certain and straightforward, they may have less value for more *exploratory* projects. Where

outcomes are clear, it may be more straightforward to work backwards to identify requirements, dependencies and assumptions.

For exploratory projects, benefits may not be clear, certainly at the outset and how to achieve them may also not yet be clear. Indeed, this has been identified by Marchand and Peppard (2013), in relation to big data projects, who call for different, more iterative approaches to such projects to address this uncertainty. It is such *exploratory* projects, seeking new insights from data, that are the focus of this research, rather than data projects aimed at *exploitation* (March: 1991), based on existing knowledge or theory.

1.1.4 The importance of human agency

While much data analytics or BI research focuses on technical implementation challenges, several researchers argue that human and social factors are important (Marchand & Peppard: 2013, Yeoh & Koronios: 2010, Hopkins, Lavalle & Balboni: 2010, Wang & Wang: 2008, Marchand et al.: 2001). For example, managers complain of information and communication overload (Hopkins et al.: 2010, Hemp: 2009). Human and social factors seem likely to be particularly important in the context of the exploratory projects in view.

This raises the question whether or not more investment in technology simply to generate more data, is necessarily the most effective approach to dealing with the information challenges and requirements business data users face. Indeed, McKinsey research (Manyika et al.: 2011) at a sector level cites the importance of talent available and data-driven orientation as important factors to enable value to be captured from big data and related analytics investment.

At an individual level, academic research of human factors often also focuses on skills required, particularly quantitative analytical skills (Davenport: 2009), and more recently ‘data scientist’ skills (Davenport: 2014). At an organisational level, academic research also calls for organisations to develop a data orientation (Marchand et al.: 2001). However, they go on to identify a general lack of research on social aspects of data initiatives within organisations.

Subsequently, related research on measuring information management practices and related capabilities finds that senior managers don't recognise sensing activities or practices as equally important to those related to collecting, organising or processing information (Kettinger & Marchand: 2011). They argue this may point to their tacit or implicit nature or the researchers not fully capturing related activities in their related constructs. Certainly, this underlines their call for more research on social aspects of data inititiatives in general, and this more socially anchored and tacit sensing dimension in particular.

1.2 Purpose of this research and personal motivation

The starting point for my research reflects my personal consulting experience. Mixed project outcomes noted often seem to reflect a corporate and managerial emphasis on deploying technology to address perceived information and data challenges. This emphasis might result in paying less attention to understanding their information and knowledge needs, as well as how data collected and produced will be exploited by business users, to create insight and realise the associated value and benefits sought.

While the increase in data clearly seems a fundamental trend and shift for IS, the (big) data investment trend seems yet another IT supplier-led response. I am concerned that it will result in similarly mixed IT investment outcomes, as I have noted more widely over my more than 20 years of IT governance consulting. However, establishing clear project and data requirements for exploratory (big) data insight projects seems even more challenging than for other IS projects.

Therefore, this research seeks to address the lack of research outlined by better understanding the social context encountered within data projects. It does so by starting with how participants frame and pursue such data initiatives and the challenges they encounter. This is grounded in the premise that these initiatives are essentially about learning from data and putting such learning into practice.

This research seeks to address the following question: **How do participants in exploratory data initiatives collectively use data to pursue insight?**

1.3 Structure of the Thesis

This research addresses the challenges outlined above by focusing at the level of projects or initiatives, rather than at the level of the individual or organisation. Projects or initiatives represent a common site or location where new insights are pursued or a vehicle for doing so. The research embraces a human starting point, adopting Wenger's (1998) communities of practice (CoP), situated learning lens, to see what this reveals about how insight and realising related benefits are pursued, that may be missed when adopting more instrumentalist or processing frames and logics. The research examines exploratory data initiatives empirically within two contrasting organisations, to better understand the participant frames, logics or theories and practices at work, and the related social challenges data project teams encounter.

As explained below, the rest of the thesis is organised into the following chapters:

- Chapter 2: Literature review and positioning the research
- Chapter 3: Research design and rationale
- Chapter 4: Introducing the cases
- Chapter 5: Details of data analysis undertaken
- Chapter 6: Thick case descriptions, as two monographs
- Chapter 7: An explanatory framework and reflexive case comparison
- Chapter 8: Discussion and theoretical contributions
- Chapter 9: Practical Implications
- Chapter 10: Validity and limitations
- Chapter 11: Conclusion

Knowledge and insight are used interchangeably within the thesis, as explained in Chapter 2. Starting with the IS field, Chapter 2 then reviews several related literatures in terms of what they have to contribute to thinking about how to generate new knowledge or insight from data. The literature review serves to locate the research and the specific research question already posed. Several important ideas are presented which emerged from the literature review and

informed my fieldwork. Finally, it introduces and examines two areas of research that emerged as important during my data analysis and writing up my discussion section, where they are reexamined in light of my field observations.

Chapter 3 explains the rationale for adopting an ethnographic research strategy, and provides a description of my approach. This recognises the recursive nature of my own research effort – trying to derive new insight from the data collected. The cases and particular data initiatives examined are then introduced in Chapter 4, before describing how related data collected was analysed in Chapter 5.

Chapter 6 presents my findings in terms of rich or thick descriptions, as two monographs, one in relation to each case, describing what was observed in relation to each case. Chapter 7 introduces an *explanatory* framework, reflecting my initial sensemaking of how insight from data is pursued within the cases examined. Using this initial explanatory framework as a starting point and structure, a cross-case comparison of the two cases and my own research is presented to more richly explain what was observed. My research is introduced as a third case in order to reflexively compare and contrast my research with the cases studied in a structured way. An enriched explanatory model is then presented as a basis for discussion.

The Chapter 8 discussion presents an argument for two main contributions, one to extend and refine current IS theory, and the other to researching data initiatives. Chapter 9 touches on the practical implications of the research, offering some preliminary governance questions for managers to pose when framing such initiatives for themselves and other participants, and to use when planning such projects. Chapter 10 reviews various validity criteria in relation to the research and recognises several limitations of the research undertaken.

The thesis concludes by calling for further research on the important social aspects of data initiatives, in particular with regard to data use. It positions the emerging explanatory framework and MSS improvements as useful starting points for such further research, as well as for practitioners to use to improve their understanding of these aspects of such projects. Several appendices are

also included to provide transparency of the research process, in particular about the data collected and how these were analysed.

1.4 First person presentation style adopted

In the light of the ethnographic research approach adopted, the first person is used extensively throughout this thesis. My rationale for adopting this stylistic convention is provided in the Method section in Chapter 3.

2 PURSUING INSIGHT FROM DATA – A LITERATURE REVIEW

Having already argued for the need for more research to better understand the human and social aspects of data initiatives, this chapter focuses on establishing what is currently known about what makes for the successful pursuit of insights from data. This fundamentally informs the research undertaken and the thesis presented, rather than providing a more traditional literature review.

After considering definitions of some of the core concepts addressed in the thesis, the chapter examines what the Information Systems (IS) field has to say about pursuing insight from data. Two main IS perspectives are identified and briefly described. A closer review and critique is then presented of the MSS model, representing the best summary and related available model identified within the field. In particular, I problematise the focus on improving decision-making as the natural objective or outcome for data initiatives, common to both IS perspectives, as too narrow and at the same time too abstract to be useful.

Several adjacent disciplines, also concerned with studying the pursuit of insight, are then considered: cognition, knowledge management, learning, sensemaking and research. Their characterisations of the phenomenon are summarised in a tabular format in Table 2-2. The table also outlines the key related considerations and the challenges identified, as well as their typical research focus and unit of analysis. A comparison of these approaches forms a useful basis for framing and locating my research. Firstly, it identifies the CoP framework as a useful sensitising lens for my subsequent fieldwork observation, data analysis and related sensemaking. Secondly, it helps position my research at the intersection of Organisational Learning (OL), Knowledge Management (KM) and IS fields. An overview of the CoP framework used, as well as several important ideas and debates that emerged from these literatures and informed the research, are then presented.

A final section introduces sociomateriality concepts and relates these to efforts to conceptualise and distinguish data, information and knowlege. These emerged during data analysis, writing and reflection as important for my argument related to disentangling data from a broader MSS IT artefact, giving it more prominence in our theorising of data initiatives and related systems. In the interests of clearer recoverability (Checkland: 1999), this was kept distinct from the literatures outlined earlier in the chapter, which informed and helped frame my research inquiry and entry to the field.

2.1 Definitional Clarity

A definition of 'insight' is not readily identifiable in the IS literature and this thesis uses it interchangeably with the more common term 'knowledge', which is explored further below. While data can take many forms and are of pervasive concern to many fields, the capture, storage, processing and use of *digital* forms of data has always been a central concern for the IS field. While a broad concept of data is used, it is this *digital* form of data that is of particular interest.

A lack of consensus and definitional clarity has long existed within the IS field about basic data, knowledge and information concepts (Kettinger & Li: 2010, Checkland & Holwell: 1998) and their interrelationships. Many of the existing definitions are grounded in what has been characterised as a dominant IS data processing perspective (Shollo & Galliers: 2013), where data are processed to produce information. Indeed, Kettinger and Li's (2010) recent contribution to this debate, the knowledge based theory of information, characterises information as a function of data and knowledge, which can be expressed as follows:

$$\text{Information} = f(\text{Data, Knowledge})$$

This is reminiscent of the instrumental logic highlighted in the introduction. It treats Data and Knowledge in terms of sets of condition-action logical pairs. This is predicated on a view of data as measurements or descriptions of attributes about events or objects, knowledge as justified true belief about the relationship between concepts, and information as the meaning produced by the application of a knowledge framework associated with a selected state of

conditional readiness for particular goal directed activities. While they argue for an interactive view of data and knowledge to produce meaning, they describe knowledge as “a set of rules produced by human societies, such as the condition-action pairs that specify a law-like relationship” (Kettinger & Li: 2010: p.412). This represents a highly codified view of information and knowledge.

Checkland and Holwell (1998) provide an alternative, human-centred view of these concepts and their interrelationship, though still treating data fairly straightforwardly: as *facts* or selected facts, which they term *capta*, about an observable world. They see this selection being influenced by cognitive factors and characterise information as meaningful facts, influenced by context and interests, while they characterise knowledge as larger, longer-living structures of such meaningful facts. Their description of the interrelationship of these concepts is presented visually in Figure 2-1.

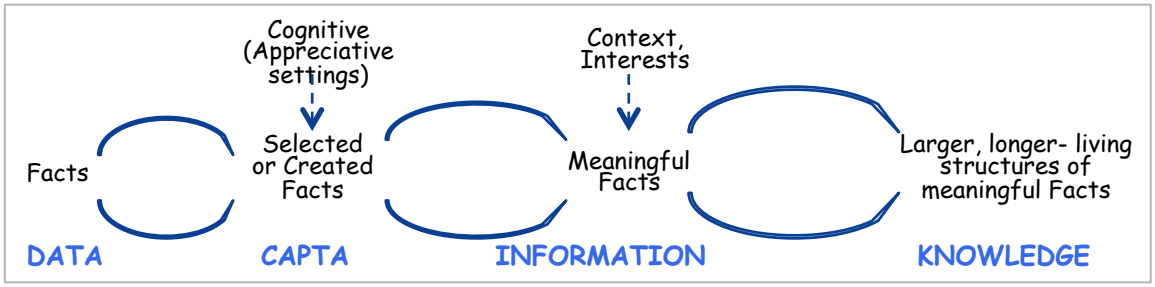


Figure 2-1 The links between data, capta, information and knowledge
(Checkland & Holwell: 1998: p.90)

While they don’t explore or describe in detail the human processes and activities involved in generating knowledge or information from data, their characterisation implies an interactive view. Their thinking provided a useful human-centric starting point for my thinking about these terms when entering the field. These concepts, related distinctions and their interrelationship are revisited in the discussion and wider implication chapters in the light of my observations and related reflection.

An important inherent part of the activity of deriving and discovering insight from data may be the use of tools or technology as an aid to capture, store and process it. Indeed, it is usually the trigger or prompt for considering this as an IS

concern. Within this thesis the term *tools* is used to mean the *software* or application elements used by project participants to facilitate the capture, organisation, processing and engagement with data, rather than the underlying technology hardware infrastructure or sensors used. This also highlights my interest in the phenomenon within a project context rather than analysts working in a permanent analyst role. My project level phenomenon is described in more detail in section 2.7.4.

A common call to improve theorising IT artefacts and their use is also noted within the IS field (Clark et al.: 2007, Orlikowski: 2006, Orlikowski & Iacona: 2001), which emerged as relevant when writing up my research. This literature is introduced and considered further in section 2.8.

2.2 IS Perspectives on Generating Insight

While generating insight from data should be a fundamental concern for the IS field, it has traditionally been addressed indirectly, under topic areas reflecting related systems or technologies involved, reflecting how data are often embedded in such systems. For instance, much of the related research identified during my literature review was found using keyword searches related to Decision Support Systems (DSS), Executive Information Systems (EIS), Business Intelligence (BI), and Knowledge Management Systems (KMS). A more general category was later identified: Management Support Systems (MSS), which includes the other system types (Clark et al.: 2007). The specific topic area *data analytics* was also identified as emerging more recently (Chen et al: 2012, Davenport et al.: 2001), often in the context of big data.

Writing in the context of BI, Shollo and Galliers (2013) identify two strands of thinking within IS, a data processing perspective and a more human-centred perspective. These perspectives are briefly outlined in turn below, before more closely examining and critiquing the MSS model. Their common focus on decision-making, as the natural outcome or objective for data and related systems use, is then problematised.

2.2.1 Data Processing or Information Communication Technology (ICT) perspective

Shollo and Galliers (2013) identify a dominant perspective, characterised by the instrumental logic outlined in the introduction. This perspective characterises generating insight as transforming data into information (or knowledge) through analysis and the application of prior knowledge, typically codified in processing rules (Kettinger & Li: 2010). This is termed the Information and Communications Technology (ICT) perspective (Shollo & Galliers: 2013) and identifies informing better decision-making as its overarching objective. It has a long history and is grounded in the related input-processing/storage-output systems model used extensively in IS teaching (Davis & Olson: 1984).

So far as this dominant view identifies human factors relevant for *implementing* MSS (which include BI systems), Clark et al. (2007) mainly consider and include these indirectly. This model is examined and critiqued more closely in section 2.2.3.

2.2.2 Human-centred IS perspectives

The dominant perspective within IS outlined above has been criticised for its lack of focus on the human *sensemaking* involved in generating insight (Weick: 1995). Shollo and Galliers (2013), identify an alternative strand of IS research focused on *organisational knowing*, which draws on the socially constructed strand of KM research (Tsoukas: 2005, Blackler: 1995). This strand acknowledges limitations when viewing knowledge as a *resource*, based purely on explicit or encoded forms of knowledge. Instead, it argues for a socially situated, constructivist view of knowledge (Newell & Galliers: 2006) and a focus on *activities* involved.

The research reviewed in this area focuses on social aspects of knowledge *sharing* and how this is facilitated or constrained through the use of tools, rather than specifically on how new knowledge *creation* or *discovery* occurs. A partial exception is a study by Shollo and Galliers (2013), focused on BI systems use, to facilitate organisational knowing. Acknowledging the role of prior participant

knowledge as a starting point for analysts, they identify two cyclical processes triggered by BI use: problem articulation and related dialogue, as well as data selection activities, arguing they are facilitated by the BI providing a solid foundation of facts and a functional capability to *drill down* and *roll-up* relevant data, often as evidentiary support for discussion.

Wang and Wang (2008), in relation to data mining, also point to the importance of learning and knowledge sharing activities. In particular they highlight the different roles played in the pursuit of knowledge discovery, by analysts on the one hand, and by business insiders on the other, arguing for interacting learning and data mining cycles.

Communities of Practice (CoP) and boundary objects

The knowing strand of IS acknowledges the importance of CoP as sites for sensemaking to occur (Newell & Galliers: 2006). As also highlighted above by Wang and Wang (2008), it recognises the likely importance of interdisciplinary interaction as a source of generating new insights. However, work on interdisciplinary data sharing, by computer scientists trying to relate fields of *domain* knowledge, points to significant definitional challenges (Ribes and Bowker: 2009). They term related codifications of definitions and relationships or mappings between different knowledge domains *ontologies*. To avoid confusion with the term's meaning within research methods (Blaikie: 2007), the notion of *economies of meaning* (Wenger: 1998: p.198) is preferred and used instead. This concept is explained further in section 2.6.

In subsequent, related research, Edwards et al. (2011), argue for the importance of such *metadata* as a vehicle for reflection and communication within and between practitioner or disciplinary groups. They also argue for the importance of related social processes rather than simply seeing metadata as fixed or as a *resource*. This concept seems closely related or equivalent to what Wenger (1998) terms a Boundary Object. Star (2010), who first introduced the idea of a Boundary Object, argues for it being an emergent, flexible response to information needs, facilitating different groups working together without the need for consensus. When these become agreed and institutionalised (i.e.

standardised or codified), she argues these should be differentiated as Boundary *Infrastructure*, calling for more research on how this transition is achieved.

Sociomateriality of IT artefacts

The knowing strand outlined is also consistent with a growing recognition of the sociomaterial nature of IT artefacts (Cecez-Kecmanovic et al.: 2014). Sociomateriality provides a more complex view of the IT artefact as a phenomenon, highlighting various important social aspects and activities that are important, However, no studies were found researching data initiatives using this perspective or seeking to theorise data as a distinct type or class of IT artefact, or element.

Therefore, neither sociomateriality, nor the social strand of BI outlined, currently provide a ready framework that can be used as a starting point for researching the social aspects of data initiatives (or for practitioners to better manage big data initiatives to achieve envisaged benefits). Indeed, the alternative, human-centred strand calls for more research adopting these perspectives, especially more practitioner-oriented empirical research (Cecez-Kecmanovic et al.: 2014).

Nevertheless, certain aspects of this literature emerged as relevant during analysis, reflection and writing up my research. Its relevance emerged in relation to my argument for disentangling data from a broader IT systems artefact or construct, as used in the MSS model, which is examined more closely in the next section. These aspects of sociomateriality are introduced and considered further in section 0.

2.2.3 A closer examination and evaluation of the MSS model

Although it has limitations, the MSS model represents the most comprehensive summary of what is known within IS currently about such systems, and provides a starting point for my research. Other models identified, such as the Information Management Practices (IMP) construct (Kettinger & Marchand: 2010) and the interacting learning cycles model (Wang &

Wang: 2008), either focus at an organisational level or on ongoing interaction and learning rather than on project contexts.

An overview of the MSS model

Clark et al. (2007) seek to identify common elements of MSS that make for successful implementation and adoption of such systems. They present a systemic, causal model, which is grounded in an extensive review of various areas of IS literature, covering systems designed to support management action and decision-making. These areas include literature on EIS, DSS, KMS and BI systems. As such, the MSS model seeks to represent a comprehensive summary of IS knowledge in this domain.

Firstly, they validate and refine their causal model for 'face validity' with 23 Chief Information Officers (CIOs) attending an IBM training conference in New Jersey, during a one-day focus group. Secondly, they build a corresponding parametric model using a systems dynamic approach and related tool (iThink) to test the logic of the model outcomes against reasonable expectations for a number of scenarios. They argue that their testing provides support that the logical structure of the model appears sound.

Table 2-1 describes the key causal constructs they identify as factors for successful MSS implementation and indicates the related response constructs they impact on. For those relationships where the authors note no empirical evidence, I have italicized the related response constructs to highlight this.

Table 2-1 MSS Success – causal and response model constructs

Causal Construct	Brief Description	Response Constructs
Executive Commitment to the MSS	Executives provide the resources necessary for an MSS and the impetus for its use. Developers working with them is seen as critical to gaining a broad base of support for an MSS.	Level of MSS Technology Desired User Commitment to the MSS
Perceived MSS Benefits	The perceived benefits that may accrue from using the MSS. These may not be realised depending on linkages to business need or objectives and organisational readiness to capture benefits. Perceived benefits, relative to costs, influences levels of commitment. Benefits relate to deeper understanding, challenging or verifying assumptions and are considered more difficult to achieve in more complex problem spaces. Other challenges or hurdles relate to leveraging tacit and unstructured forms of data.	Executive Commitment to the MSS User Commitment to the MSS
Management Decision Quality	Improved decision-making is seen as the objective for MSS Use, traditionally in relation to critical success factors for an organisation. However, decision effectiveness is recognised as multi-dimensional and difficult to judge or measure. The authors call for more research on MSS goals.	Perceived MSS Benefits
User involvement in development	User involvement in DSS is characterised as active and iterative, between users, the MSS and analysts. Though seen as an important factor for user MSS satisfaction, mixed results are noted in relation to decision quality. Related benefits may depend on the MSS, the nature of the development process, and communication and alignment between users and analysts. Where the MSS involves significant knowledge acquisition across a number of areas, user involvement is considered critical.	MSS to problem space match *MSS costs* *User's MSS knowledge base required*
User Commitment to the MSS	This is distinct from executive commitment and influenced by different factors. Resistence to knowledge sharing is noted as an example of a lack of user commitment. User commitment is seen as influenced by user knowledge of the MSS and the wider problem space being addressed, as well as perceptions of personal benefits and usability of the MSS.	User Involvement in Development *Level of MSS Technology Desired* Level of MSS Use
MSS costs	Organisations typically encompass hardware, software and personnel costs, though with less focus on user training in development costs. Costs are considered a key constraint to building MSS and related benefits. Unstructured projects are recognised as difficult to cost and traditional approaches as ill-suited. The view adopted is one of investment in managing an organisation's knowledge resources.	*Executive commitment to the MSS* *User commitment to the MSS*

Causal Construct	Brief Description	Response Constructs
Usability of the MSS	Usability factors include the complexity of the system, problem space, type of technology, as well as user ability and knowledge. This is considered different for power users or information providers versus other users or information consumers, with different tool requirements. Far fewer of the former are found than the latter, yet most tools are designed with the former in mind. The key consideration is eliciting the right information from the MSS. Traditionally this has focused on specifying appropriate models rather than their validity, usability or the ability to integrate them with more tacit information and experience, which have received less attention within decision-making.	User commitment to the MSS MSS quality
MSS functionality	This is defined as the extent to which a system supports the requirements of the decision-maker using it. It is comprised of various dimensions, going beyond features provided. One dimension is supporting various stages of decision-making: intelligence, design, choice, implementation and feedback. It must also be flexible enough to facilitate exploring the problem space and generate alternatives, and rapidly modifiable to allow for changing the problem space and extending its structure, or developing new applications. This recognises that users cannot anticipate their requirements until they see what is possible. It also anticipates the need to support changes in the problem space over time, though also noting stabilisation of the problem space with routine MSS Use.	MSS Problem Space match MSS quality
User's MSS knowledge base	This is defined at the organisational level as the accumulation of knowledge and experience through ongoing, iterative use of the MSS to address a particular problem space. This is primarily in relation to identifying new relationships between key variables or recognising fallacious logic, increasing their insight into the problem space or domain, characterised as a set of action-outcome relationships. It recognises an individual starting point but encompasses how this is shared, integrated, absorbed and contextualised more widely. This is recognised as complex and ongoing.	User commitment to the MSS Usability of the MSS User's MSS knowledge base required
MSS Training	This is defined as the organised instruction users receive in relation to the MSS, encompassing the decision structure, as well as the MSS technology (software and hardware). Shortcomings in training are highlighted in relation to user understanding of underpinning data models within MSS, inherent limits to pre-implementation training in the light of complexity and minimal training in general.	User's MSS knowledge base required MSS costs
User's MSS knowledge base required	This is not clearly defined but seems to recognise that the actual, current knowledge of users about the MSS and related problem space may fall short of that required to make effective use of the MSS.	MSS training

Causal Construct	Brief Description	Response Constructs
Problem Space Match	This is the degree to which the MSS addresses the problem space and is indicated as one of the major factors determining use of the MSS and MSS Quality. Systems design and development is seen as key to achieving this, though also a key bottleneck, given the considerable challenges noted. These include determining the relevance of a growing volume and variety of data available, encompassing a variety of users and dynamic problem spaces. Iterative approaches, focusing on data use in decisions, and integration with KMS are noted as recommendations to improve this.	Perceived MSS Benefits MSS to Problem Space Match Level of MSS Use Management Decision Quality User's MSS Knowledge Base Required
MSS to Problem Space Match	It is not clear how this construct differs from the problem space match, though both terms were distinguished in a relationship table, with this representing a response construct of problem space match. This highlights that the construct of a problem space is not clearly defined by the authors.	MSS Quality
Level of MSS Use	The amount or level of use of an MSS over time is highly correlated with decision quality and effectiveness. However, given conflicting research results, the authors call for further research on the impact of MSS Use on decision effectiveness. Factors influencing use are identified as the accuracy and relevance of MSS output as well as the user's knowledge of both the problem space and how to use the MSS. The culture of the organisation in which its use is being enacted is seen as part of the problem space and its complexity.	User's MSS Knowledge Base MSS Costs Management Decision Quality
Technology base	A wide variety of tools and technologies are recognised as relevant and used by organisations to build MSS solutions, acknowledging an implicit assumption that tools are used. Its boundary is seen as the sum of all the hardware and software devoted to the MSS. The technology base used is seen as a balance between technology requirements and resource constraints and the ability of the organisation to select and deploy appropriate tools is seen as critical to adoption.	Technology gap MSS functionality Usability of the MSS MSS Costs User's MSS knowledge base required
Development effort	While not specifically defined, this is clearly seen as part of the resource requirements and potential constraints resulting in the technology base deployed in an MSS.	MSS technology base
Level of MSS Technology Desired	This is not specifically defined, though clearly related to MSS functionality requirements and subject to resource constraints, as well as the technology available.	Development Effort
Technology Available	This is seen as the technology available to an organisation to acquire and deploy within the market in which it operates.	Technology Gap Development Effort
Technology Gap	This is seen as the difference between the present MSS technology deployed within an organisation versus what is available to it, to address its MSS technology requirements. This is primarily seen as driven by resource constraints, though also more broadly related to the organisation's absorptive capacity (Cohen & Levinthal: 1990) to recognise, assimilate and apply new knowledge. In this instance the effective use of relevant technology.	Executive Commitment to the MSS

Causal Construct	Brief Description	Response Constructs
MSS Quality	This is described as the overall ability of the MSS to provide efficient and effective support to the user during decision-making. It is also seen as critically dependent on the MSS to Problem Space Match. They recognise that factors determining MSS Quality have not been fully identified, especially in relation to what gives rise to effective use, citing mixed outcomes from MSS efforts. The authors call for more research in this area.	Management Decision Quality

Clark et al. (2007) then describe a model based on five structural characteristics which define systems, drawn from systems theory, to which they align particular constructs:

1. **Objectives:**
    a. Perceived MSS Benefits
    b. Management Decision Quality
    c. Overall MSS Quality
    d. MSS to Problem Space Match
2. **Environment:**
    a. Problem Space Complexity
3. **Resources issues:**
    a. MSS Technology Base
    b. MSS Costs
    c. User Involvement in MSS Development
    d. MSS Training available
4. **Components:**
    a. MSS Functionality
    b. MSS Usability
    c. MSS User's Knowledge Base
5. **Management:**
    a. Executive Commitment to the MSS
    b. User Commitment to the MSS
    c. Level of MSS Use

Based on their review of the literature in relation to the constructs outlined, the authors argue that behavioural issues in connection with MSS, while critical, have been widely addressed. By contrast, they argue that MSS design aspects are under-researched, especially MSS functionality and MSS technology. They also note little research on MSS Cost.

A diagram of the MSS model and relationships they posit between constructs is depicted below. The causal model depicted in Figure 2 2 provides a visual summary of existing IS research about factors that are important to successfully implementing MSS and their interrelationships. Those constructs that seem related to human attributes or activities are highlighted in bold red text. The authors recognise measurement challenges for several constructs, and indicate various relationships that are suggested but with limited or no empirical evidence (as highlighted in Table 2-1), and call for more research in a number of areas.

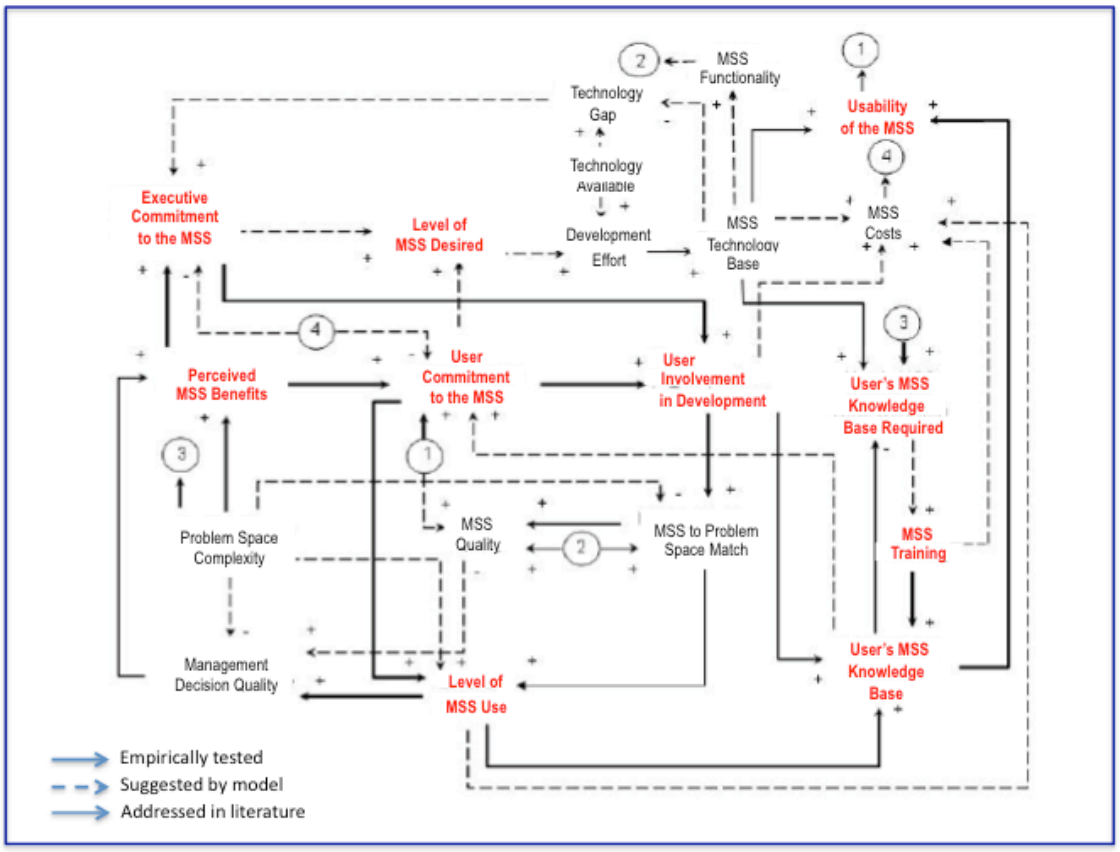


Figure 2-2 Dynamic Structure of MSS (based on Clark et al.: 2007: p.588)

General Critique

As can be seen from Figure 2-2, the authors seek to avoid reductionism by identifying various complex interaction effects between causal and response constructs. Nevertheless, they label and define some of these constructs quite narrowly, e.g. User MSS Knowledge Base and MSS Training don't clearly

encompass broader knowledge that they acknowledge elsewhere as important, e.g. in relation to the Problem Space rather than just the decision process. They also fail to clearly define all constructs used in their model, leaving some ambiguity, e.g. MSS to Problem Space Match, with further examples highlighted in Table 2-1. However, this may well reflect article length constraints.

The different ontological and epistemological assumptions and perspectives inherent in the different studies they draw on and seek to synthesise are not addressed, nor do they make their own stance explicit. This makes it difficult to clearly position their study in terms of ontological and epistemological perspectives. However, their use of a hard systems model and focus on causal relationships implies a realist ontology. Their epistemological assumptions are more difficult to identify. The level that the model is focusing on is also not made clear, identifying both individual level factors, such as commitment, as well as environmental factors in their model.

Critque of the model for my research purpose

The most notable feature of the MSS model is its central focus on the MSS tool or system, in particular effective implementation and adoption in order to improve management decision-making. MSS Use and design seem secondary foci. We could summarise the underlying logic of this model as follows:

Users + **Improved** MSS Tool (implementation) = **Improved** Quality Decisions

By contrast to this tool implementation focus, my research has a fundamentally different starting point – a focus on the broader human or practitioner endeavour, i.e. generating insight about a phenomenon of interest, rather than implementing an MSS to improve decision-making. The MSS focus on improved decision-making is common across both IS perspectives outlined earlier and is problematised separately in the following section, given its prominence within the field.

Another distinction is that while Clark et al. (2007) seek to address all MSS, my research focuses only on a specific subset of MSS. In particular, my research focuses on those MSS used in *exploratory* data initiatives in pursuit of new

knowledge, rather than data initiatives that seek to *exploit* data using existing knowledge. The latter is likely to be a common and quite valid objective for a much wider set of MSS. Therefore, my critque below relates specifically to this narrower set of MSS.

In terms of pursuing new insight, the MSS model doesn't explicitly address the human agency and sensemaking involved. Instead, it offers a more technical description of an MSS and its related implementation. This may reflect the systems model they use, in terms of which they only explicitly identify human agency under the related management heading, in terms of commitment and level of MSS Use. However, human factors are also inferred when describing the following MSS model constructs: problem space complexity, a system or tool's match to this, tool and problem space knowledge and training, and indeed as antecedents to system or tool use and commitment too.

The result is that the human actors involved are somewhat lost from view, especially in terms of being able to clearly link factors to particular human characteristics and activities. For instance, there seems to be no mention of IS practitioners within the model, which points to the likely incompleteness of the model described. It also makes it difficult to clearly isolate and distinguish between individual, social, technical and context related factors, which would enable clearer measurements and related fieldwork.

The MSS view of technology also seems unitary, viewing it as a black box, though comments in relation to particular constructs imply the importance of particular elements, for example, the characteristics of its underlying data model. While they acknowledge an alternative, more interactive or dynamic relationship between technology and actors, citing Orlikowski (1992), their model nevertheless seems to reflect a technology adoption view. Such approaches have been criticised by Orlikowski and Iacona (2001) as resulting in under-theorising the IT artefact.

I will return to the MSS model in the discussion section, to review it further in the light of the research findings emerging from my cases and cross-case analysis.

2.2.4 A narrow focus on tools and improved decision-making

A prominent feature of the MSS model, and indeed of both strands of IS thinking about generating insights, is a broad acceptance that the objective or desired outcome and benefit of doing so is *improved decision-making* (Shollo & Galliers: 2013, Clark et al.: 2007). From the outset, this struck me as both a very narrow, and at the same time a very abstract framing of potential benefits. Indeed, this may represent a form of institutionalised thinking within the IS field.

On the one hand this focus on decision-making doesn't appear to allow for the readily forseeable outcome of increased understanding or knowledge (i.e. learning). Improved knowledge and understanding about a phenomenon *may* in turn lead to more informed and sophisticated interpretation or sensemaking about the phenomenon, and more informed action (or indeed inaction) in relation to it. This may implicitly or explicitly involve decision-making but implies an intermediate outcome or enhanced *capacity*, rather than a straightforward, immediate link or path to decision-making or action.

On the other hand, *improved decision-making* is somehow too vague and abstract to be useful to readily appreciate the likely benefits that might arise. Instead, improved decision-making needs more specificity to be meaningful or useful. In particular, it needs to be understood in relation to a particular phenomenon and situation. However, while the importance of such context is recognised, it is typically referred to as the 'problem space' (Clark et al.: 2007). This frames the activity as problem-solving, when pursuing insight may also seek enhanced understanding or be more opportunistic and exploratory in nature. In other words, it can encompass both exploitative and exploratory objectives for data use (March: 1991). Eisner (2003) also highlights that adopting different search strategies may throw up very different challenges. Opportunistic searches may require more creative and sensemaking approaches to deal with considerable ambiguity, while problemistic searches seem to lend themselves to more straightforward benchmarking, root cause analysis and traditional analytical approaches.

While the above points to a broader and richer view of outcomes from the pursuit of insight, improved decision-making clearly remains a valid objective and outcome. However, such decision-making may be intertwined within action and may remain implicit or tacit, rather than being readily identified and ascribed to improved knowledge by participants. The instrumental or information processing approach to generating insight described is also underpinned by a fairly narrow, rational view of decision-making, which has long been problematised. It has been argued that this view does not reflect actual managerial decision-making practice (Lindblom: 1959). Even managerial *information use* (what I would really term *data* use) has been found to be pragmatic and partial in the face of ambiguity (March: 1987) or limited managerial time (Mintzberg: 1975). Indeed cognitive biases have long been recognised as impacting on data and information use within IS (March: 1987, Davis & Olson: 1984), which are in turn likely to impact on data-led decision-making.

A greater abundance of available data may increase the scope and opportunity to apply analysis and adopt rational decision-making approaches (Davenport: 2009). However, adoption of new analytical approaches is hardly assured (Fekri: 2010) and is still likely to be more suited to certain, bounded or structured problems than to complex, unstructured ones (Lindblom: 1959, Davis & Olson: 1984). Snowden and Boone (2007) draw attention to context as important for decision-making, arguing that successful leaders adopt a situational approach. They identify four broad contexts: simple, complicated, complex and chaotic, recommending the following respective leadership decision-making approaches for each: categorise, analyse, probe and act.

Snowden and Boone's (2007) ideas of probing and acting seem to align with Weick's (1995) ideas of enactment during sensemaking. In particular, their ideas seem to connect with Weick's argument that while more information can be helpful to address *uncertainty,* it does not help address *ambiguity*, which requires human sensemaking. See Appendix E.2 for an explanation of the distinction he draws between ambiguity and uncertainty. Its importance here is

that ambiguity may well be a prominent feature of exploratory data initiatives aimed at deriving and discovering new knowledge. Certainly, some evidence suggests policy makers struggle to formulate new policy in the face of too much data (Dunlop: 2009). This recalls Kettinger and Marchand's (2011) finding that managers don't fully recognise what they term *sensing* practices, i.e. identifying relevant data and information. The question of information overload has also been well documented, though, intriguingly, organisations that have more data analytics experience find this less challenging (Hopkins et al.: 2010), implying a better ability to filter and select relevant data.

With these reservations in mind, I sought to avoid a narrow focus on decision-making as the natural ultimate benefit or outcome, instead keeping an open mind about additional learning and action outcomes. I also sought to take careful account of the wider context of data initiatives.

2.3 Wider perspectives on pursuing Insight

Having reviewed the IS literature, several adjacent disciplines, also concerned with studying the pursuit of insight, are now considered. The objective is to identify what they might add to IS thinking about pursuing insight within projects, as a potential starting point for framing my research.

These disciplines or literatures are highlighted visually in Figure 2-3. A tabular summary is also provided in Table 2-2, briefly outlining their characterisations of pursuing insight, key ideas and problem or research focus. This provides a useful basis for comparing them, locating or positioning my research and explaining my subsequent use of particular ideas. In particular, this is used to explain the adoption of the CoP framework for my research, providing an overview of the framework and its principal ideas. The rest of the section outlines other key ideas and debates emerging from these perspectives that informed my research.

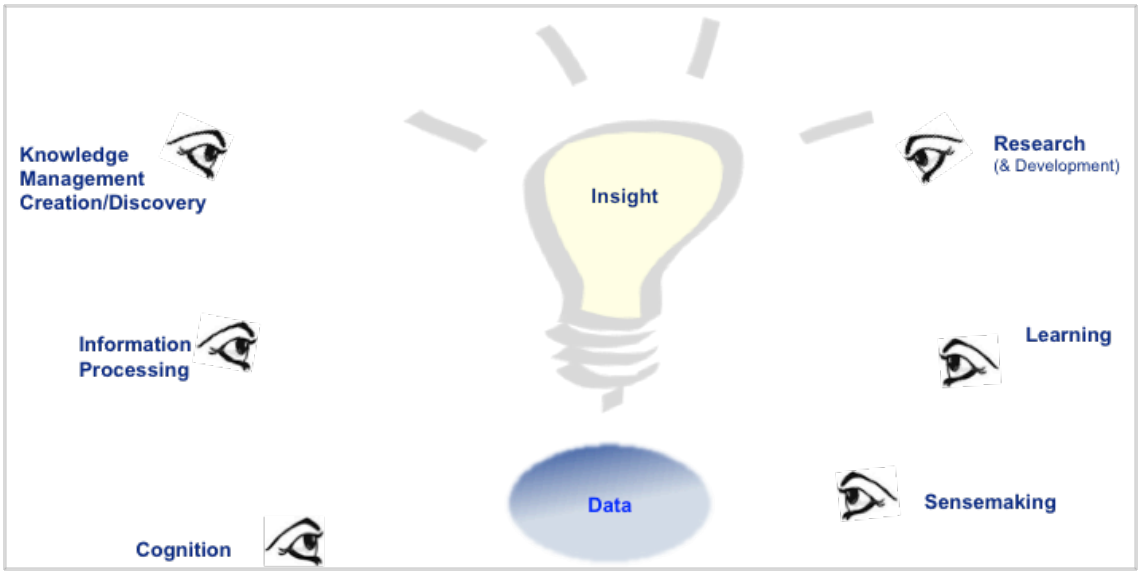


Figure 2-3 Different characterisations of how insight is pursued or generated

Both the information processing and human IS strands outlined recognise the importance of either knowledge or *knowing*, pointing to a clear overlap with KM theorising and concepts. Closely related to knowledge, learning loomed large, providing a potential process or mechanism for how new knowledge is created or discovered. Cognition, sensemaking and research fields were also identified as likely to be relevant and similarly focused either on interpretation or new knowledge discovery.

2.3.1 Various characterisations of pursuing insight

The characterisations of deriving or discovering insight provided by each field are summarised in Table 2-2 on the following page. Critical elements and challenges they highlight are noted, which are relevant to studying the phenomenon in a project team context, together with key references drawn on, their typical unit of analysis and research foci.

Table 2-2 Various characterisations of knowledge creation – critical elements and challenges (based on Douglas & Peppard: 2013)

Characterisation	Critical elements and Challenges	Unit of Analysis/Focus
Information Processing/ICT Transformation of data into information (or knowledge) through analysis and the application of prior knowledge and techniques, typically in the form of codified processing rules or algorithms (Kettinger & Li: 2010). Focus on informing better decision-making (Davenport: 2009).	Stresses the importance of problem space complexity, how a system or tool matches this, the use and commitment to systems or tools, as well as tool knowledge and training as important factors which can be identified as human related (Clark et al.: 2007).	Organisational, project level Systems, tools, algorithmic techniques Design, implementation, adoption, use of systems, project/organisation outcomes
Cognition New tacit knowledge is episodic in nature, based on accumulated experience, representing the sense that is made of current activity and experience by relating it to prior episodes or instances. At a group level, this results from active collaboration by individuals (D'Eredita & Barreto: 2006).	Processing external stimuli and data in relation to pre-existing mental models of reality. Conscious and unconscious processing is recognised with important bias and anchoring features (Davis and Olson: 1984). Importance of prior experience, memory and attention. Collaboration and shared experience important for shared cognition to emerge at a group level (D'Eredita & Barreto: 2006).	Individual level Interpretation
Knowledge Resource View Seen as social in nature and having an action orientation. New knowledge emerges and is made explicit through the interaction of tacit and explicit knowledge in a spiral model with four characteristic interaction patterns: Socialisation, Combination, Externalisation and Internalisation (Nonaka, Toyama & Konno: 2000; Nonaka: 1994).	Making tacit knowledge explicit is inherently challenging. Knowledge is situated within a specific action context and 'interaction community' or community of practice. Reflection and sensemaking activities are important, as are dialogue, language and metaphor for collective learning, sensemaking and dissemination to occur (Nonaka et al.: 2000; Nonaka: 1994).	Organisational level Codified knowledge artefacts Converting tacit knowledge to explicit artefacts, sharing knowledge
Knowing View Embodied activity, seen as mediated, situated, provisional, pragmatic and contested (Blackler: 1995). Subsidiary particulars are interiorised over time into an 'unarticulated background' which influences and frames action. Increasingly fine distinctions about a phenomenon emerge and constantly evolve via productive dialogue, reflection and action embedded in a particular context (Tsoukas: 2009, 2005).	All propositional knowledge has tacit or implied predicates. Paradoxes result from privileging abstract propositional knowledge without acknowledging their bounded time and action context, recursive nature, and the role and motives or purposes of the narrator. The importance of epistemological clarity The role of boundary artefacts as aids and reference points for cross-discipline dialogue and sensemaking (Tsoukas: 2009).	Organisational level Social and processual focus, conceptual clarity and theorising knowledge

Characterisation	Critical elements and Challenges	Unit of Analysis/Focus
Situated Learning Seen as equivalent to social, practice-based and cultural learning and distinct from individual learning focused on acquiring existing abstract knowledge. It emphasises improvisation, informality, collective action, conversation, sensemaking as well as its provisional, distributed nature (Elkjaer: 2003). New knowledge emerges through mutually constitutive processes of reification and participation, to negotiate economies of meaning (Wenger: 1998).	Engagement, Imagination and Alignment highlighted as important for emergent learning and knowledge creation. Ubiquitous and bound up with identity formation, with community membership and ownership of economies of meaning seen as important. Boundary spanning activity and peripheral participation important to encourage learning and translation, although introduces possible tensions within and between particular groups too (Wenger: 1998).	Situated individual, groups (e.g. communities of practice) Emergent, practical learning
Sensemaking Emphasises how meaning is *enacted* in a particular context, based on *cues*, especially during problem space definition when multiple approaches, options or courses of action are possible (Weick: 1995).	Resolving ambiguity is distinct from addressing uncertainty - more (appropriate) data may help with the latter but can make addressing ambiguity more difficult. Focal narrowing, to concentrate on urgent and immediate cues under extreme pressure, reduces sensemaking capacity, i.e. bounding or anchoring attention (Weick: 1995).	Situated individual, groups Interpretation, framing and sensemaking in an action context
Research Progressive clarification of, and learning about a phenomenon through structured inquiry, data collection, analysis, critical reflection and reflexiveness, enabling epistemologically grounded and appropriately qualified knowledge claims about the phenomenon in question (Blaikie: 2010).	Importance of question clarity (What, Why, How) Epistemological and Validity considerations Research design (Blaikie: 2010). Path dependency of acquiring, integrating and applying new knowledge based on a level of requisite prior understanding, a key part of Absorptive Capacity (Cohen & Levinthal: 1990).	Varied units of analysis Varied phenomena of interest Various research methods and approaches *R&D literature:* *- Organisational level* *- Industrial product development focus*

2.4 Comparing wider perspectives on generating insight

While these fields characterise deriving and discovering new knowledge differently, reflecting their particular research foci and different ontological and epistemological starting points, they all emphasise the importance of prior knowledge and some form of framing. For example, they variously characterise this framing as action-orientation, problem-space definition, research design, contextualisation, or paying attention to (or noticing) particular cues as important. Indeed, prior knowledge or experience and framing are often seen as closely interrelated.

Easterby-Smith and Lyles (2003) identify several perspectives across learning and KM fields with broadly consistent epistemological assumptions. These include an information processing perspective, which seems consistent with the historically dominant perspective identified within IS, as well as a socially constructed perspective, which aligns well with the knowing and sociomaterial strands within IS. These have therefore been grouped together in Table 2-2. Similar ontological tensions to those within IS are discernible within the learning and KM fields. These centre around the objective, codifiable nature of knowledge versus a more fluid, socially constructed view, emphasising its provisional, evolving and negotiated nature (Tsoukas: 2005, Elkjaer: 2003, Orlikowski: 2002).

Considering the fields examined in terms of their research foci and unit of analysis, highlights the particular relevance of situated learning and sensemaking literatures for examining the phenomenon at a project level. This is illustrated in Figure 2-4, enriching the characterisations depicted in Figure 2-3, to show each field's typical unit of focus and analysis (as well as some of the pertinent ideas they contribute, shown in blue italics). By contrast to situated learning, the other fields are primarily focused either at the level of the individual (e.g. cognition) or at the organisational level (e.g. research and development).

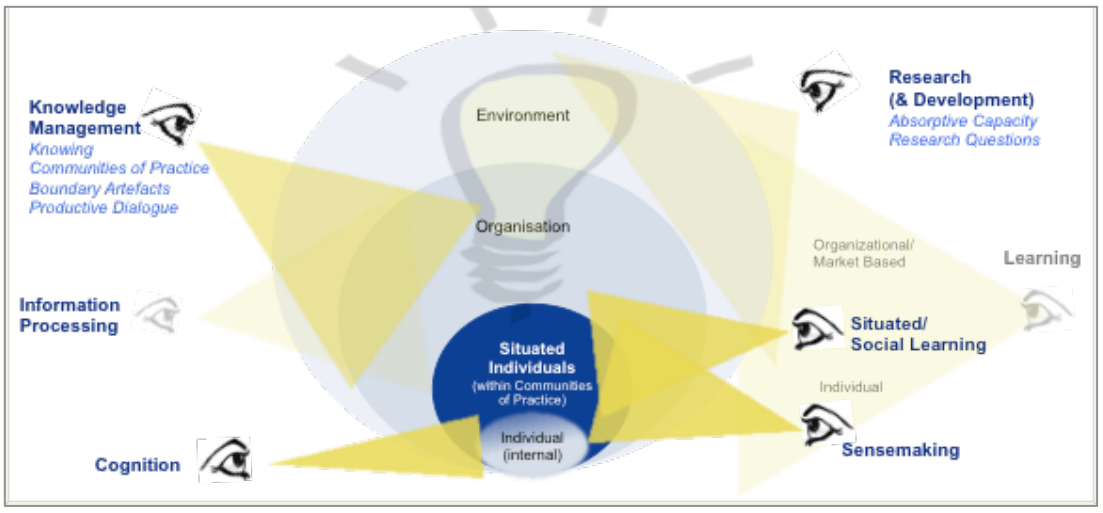


Figure 2-4 Various disciplinary perspectives on generating insights (Douglas & Peppard: 2013)

Given the human and social starting point for this thesis, and the focus on initiatives or projects, adopting ideas from the socially constructed strands of IS, KM (*knowing* strand) and situated learning were identified as particularly promising, given the consistent epistemological assumptions, as well as similar concepts and qualitative research approaches used.

2.5 Positioning the Research and Research Question

2.5.1 Intersection of three fields

Based on the literature reviewed, my research inquiry can be located at the intersection of three significant fields. While the starting point for the literature review was the IS field, the likely relevance of KM, learning (and related research) fields was quickly realised, especially as the processual aspects came more clearly into focus. Vera and Crossan's (2003) analysis of promising overlaps between the OL and KM fields proved a particularly useful starting point for locating my research in relation to situated learning within CoP.

My research represents a study of a special case of the more general phenomenon of situated learning with the CoP framework in the foreground. Its distinguishing characteristic is the focus on learning *from* or *using data*, as a central *source* or *prompt* for new knowledge and insight. This particular case of

situated learning has not been a focus for learning and KM disciplines. Instead IT is seen purely as tools or enablers (Alavi & Tiwana: 2003) as summarised in Table 2-3.

Table 2-3 IT Tools for support of Knowledge Management processes (Alavi & Tiwana: 2003: p.115)

	Knowledge Management Processes			
	Creation	**Storage & Retrieval**	**Transfer**	**Application**
IT Tools	E-learning	Data warehousing and data mining	Communication support systems	Expert Systems
	Collaboration support systems	Repositories	Enterprise information portals	Decision support systems

This focus on data also provides the main link to the IS field, bringing into view IS work on the sociomaterial nature of IT artefacts, as well as work on MSS, in particular BI and data analytics systems.

As the CoP framework provides such a central sensitising lens for my research, a brief overview of the framework is provided in the next section. At a slight distance, although still relevant, are the fields of research and development, in terms of absorptive capacity (Cohen & Levinthal: 1990), as well as sensemaking (Weick: 1995), which were not identified in Vera and Crossan's (2003) analysis. These ideas are outlined in section 2.7. Similarly, IS work on information orientation (Kettinger & Marchand: 2011; Marchand, Kettinger & Rollins: 2001), as well as work on theorising data, information and knowledge concepts, by Kettinger & Li: 2010, though undertaken from a different ontological starting point. At a more general level within IS, the (soft) systems work by Checkland and Holwell (1998) and their assessment of core concepts within the IS field are also relevant. This is summarised diagrammatically in Figure 2-5. The original elements in the Vera and Crossan (2003) diagram are indicated in *green italic* text below, while those areas added are indicated in **blue bold** text.

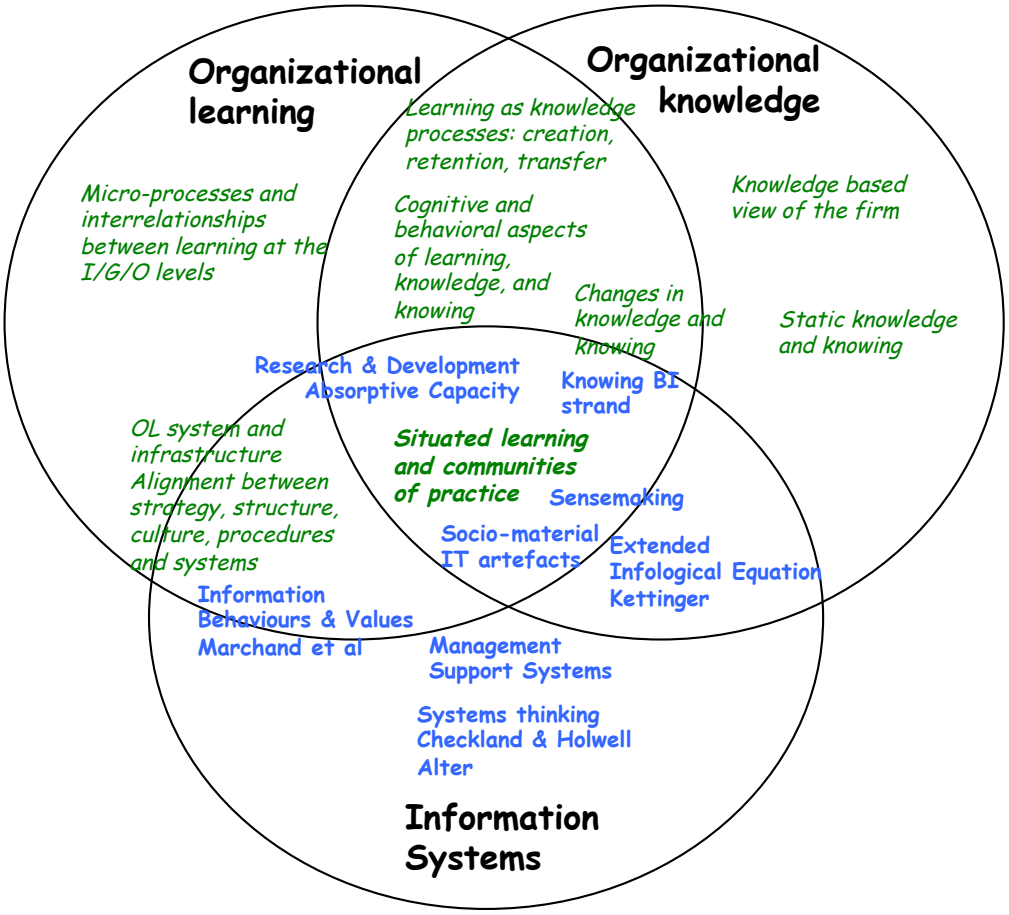


Figure 2-5 Intersection of Organizational Learning, Organizational Knowledge and Information Systems fields in relation to Exploratory Data Initiatives (extending Vera & Crossan: 2003: p.127)

As already highlighted in the introduction and above, there is a lack of research identified in relation to pursuing insight within data initiatives. Figure 2-5 helps locate this gap as well as the related research question already posed in the introduction:

How do participants in exploratory data initiatives collectively use data to pursue insight?

2.6 An overview of the Communities of Practice Framework

Given my adoption of the CoP (Wenger: 1998) as a theoretical lens with which to address this research question, a brief overview of the framework is provided, outlining its key assumptions and central ideas. The framework's usefulness for my research is then outlined.

Figure 2-6 Components of a social theory of learning (Wenger: 1998: p.5)

Wenger (1998: p.5) presents a framework comprising the following components as a basis for his social theory of learning:

- **Meaning** – "as a way of talking about our evolving ability to draw meaning from our experiences"
- **Practice** – "a way of talking about the shared historical and social resources, frameworks, and perspectives that can sustain mutual engagement in action"
- **Community** – "a way of talking about the social configurations in which our enterprises are defined as worth pursuing and our participation is recognizable as competence"
- **Identity** – "a way of talking about how learning changes who we are and creates personal histories of becoming in the context of our communities"

He grounds the framework in the following starting premises:

- Our social nature is central to our learning
- Knowledge is about competence in relation to purposes that are important to us
- Knowing involves participating and engaging in the world to pursue such purposes
- Meaning is the ultimate result gained from our learning

CoP's relevance for my research

Knowledge Management and the knowing strand within IS had already highlighted *communities of practice* as an important context for socially situated learning or generating new knowledge. Wenger's (1998) framework offers a broad conceptual framework for understanding and analysing situated learning as a process of social participation within such practice communities. He considers dialogical interaction central to such learning. He also argues that the degree to which a practice community is reflective about its practice is an important characteristic, which varies for different communities. He sees this as determining the kind of learning it engages in.

Reification and Participation

The ultimate product of learning is seen as meaning, which is contextual and located in a process of *negotiation* within a community of practice. Importantly, Wenger introduces and argues that it involves the interaction of two constituent processes (a complementary duality): 'reification' and 'participation'. His view of this concept and its duality is illustrated in Figure 2-7. A more detailed exploration of this central idea is provided in Appendix E.1.

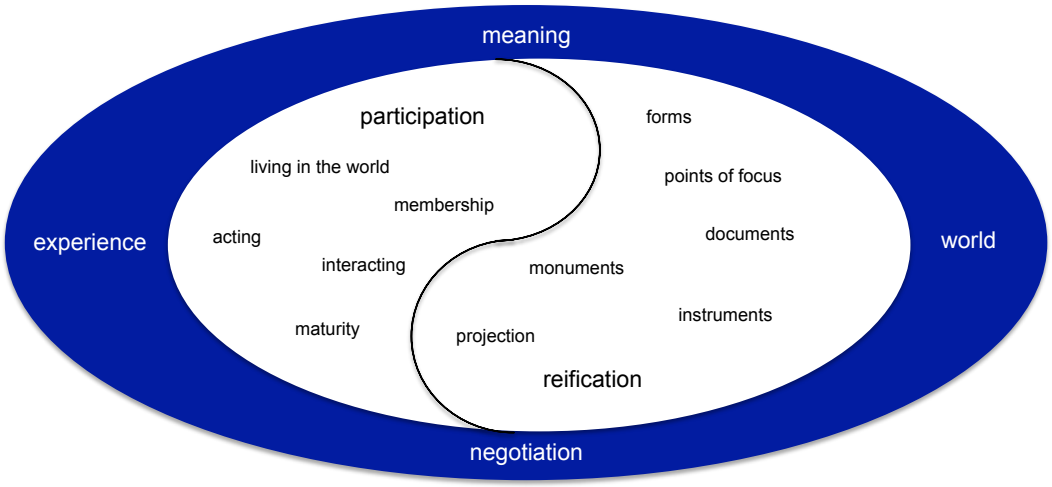


Figure 2-7 The duality of Participation and Reification (Wenger: 1998: p.63)

This characterisation echoes the KM debate about the interaction and interdependency of tacit and explicit knowledge that Tsoukas (2005) argues for. Wenger acknowledges the tacit implied knowledge intrinsic to reifications as well as the tacit dimension of practical knowledge.

Economies of Meaning

Wenger (1998) stresses the importance of identity in negotiating such meaning within a practice community's more formal structural elements, through membership. He explains how this leads to *economies of meaning* evolving through 'ownership' of meaning, recognising related power and institutionalisation. He explains this concept in terms of the generation of various meanings within a particular context, which have a different value or currency, emphasising the following characteristics:

1. "a social system of relative values
2. the negotiated character of these relative values
3. the possibility of accumulating 'ownership of meaning'
4. the constant possibility of such positions being contested
5. systems of legitimation that to some extent regulate processes of negotiation" (Wenger: 1998: p.199)

The related idea of 'ownership' of meaning recognises differences in their 'currency', different participants' levels of control over establishing and changing

such community meanings, and meaning's contested nature. Based on these ideas he argues for the importance of three processes for both identity formation and negotiating meaning: 'engagement, imagination and alignment'. He argues that they are also important considerations when formulating a design to facilitate emergent learning, acting as infrastructures for such learning.

Criticism of the CoP Framework

The main area of focus for criticism of the framework to date has been in relation to its focus on social aspects of learning, especially participation, and the potential lack of emphasis on individual learning (Hughes, Jewson & Unwin: 2007). This reflects a much wider debate within OL outlined briefly below. Indeed, Easterby-Smith and Lyles (2003) identify four different learning perspectives, including a social constructionist or social learning perspective, and incidentally also recognising clear overlaps between KM and organisational learning.

Individual versus social learning

Elkjaer (2003), in her related review of the learning field, contrasts social learning theory with *individual* learning theory, which she argues emphasises the enhancement of individual cognitive frames and privileges abstract knowledge acquisition over that emerging from practice. She cites conceptual *bodies of knowledge* as an example of such abstract knowledge, which seems very similar or equivalent to what Ribes and Bowker (2009) term codified *domain* knowledge. She sees social learning theory's starting point as our everyday lived experience and equates *social learning theory* to several other terms: *situated learning*, *practice-based learning* and *learning as cultural processes*.

She describes social learning as ubiquitous and integral to human activity, and related to the purpose of becoming a practitioner (with its associated emphasis on identity formation and the influence of social and related power structures). She goes on to characterise what it is and how it occurs as follows:

"a social learning theory emphasizes informality, improvisation, collective action, conversation and sense making, and learning is of a distributed and provisional nature" (Elkjaer: 2003: p.44)

Elkjaer (2003) characterises the tensions between individual and social learning approaches as follows: where social learning theory argues for taking a more situated or contextual approach, individual learning theory emphasises the knowledgeable, mobile individual. She also discerns two very different aims for learning – the first, a purposeful acquisition of explicit, abstract knowledge, whereas the second focuses on acquiring practitioner skills and gaining identity. For her, the aim of social learning is also less about acquiring existing knowledge and addressing known or explicitly defined problems, and more about addressing both unknown issues and what she terms 'mystery'. This seems very close or equivalent to the exploratory knowledge discovery envisaged as my research phenomenon of interest. That said, Hughes et al. (2007) argue that most actual CoP use by practitioners and research to date has not really addressed generating new knowledge. Instead, they argue it has focused mainly on transferring existing community knowledge to new joiners or 'apprentices' and emphasis participation. Elkjaer (2003) goes on to argue that people, self-evidently, engage in both types of learning, while also arguing persuasively for a synthesis of the two approaches. She turns to Dewey (1938, 1910) and his ideas of 'inquiry', 'reflection' and 'experience' as a route to such a synthesis, thereby addressing the inseparability of identity, practice and knowledge (abstract and tacit).

Other criticism

Hughes et al. (2007) highlight critics that would prefer a clearer definition or delineation of what represents a CoP, though also recognise a clearer delineation may prove problematic. Instead they focus on the need to also address constellations of CoPs and the learning that may occur between them, which they argue the framework doesn't really address.

Indeed, they argue that boundary related concepts need further development and are not clear enough. They are also concerned about the institutionalisation

of knowledge that may occur, related power dynamics and conflicts that may arise. Here, they feel that while the framework does acknowledge these aspects, it does not adequately address them. These are themes that do arise in the cases observed and are returned to in my Chapter 7 case observations and in my Chapter 8 discussion.

2.7 Other key ideas and debates that informed my research

While the CoP Framework outlined above represented the most useful lens with which to consider my cases, several other useful ideas and debates emerged and informed my fieldwork from the different literatures considered, as per Figure 2-4. These literatures are each vast in their own right and my review of them necessarily broad and partial. It focused on those elements that seemed most relevant to my particular phenomenon and a project context. It was further filtered in terms of considering those areas of each literature that seemed ontologically and epistemologically consistent with each other and with my socially constructionist ontology, outlined further in Chapter 3.

The key debates and ideas that emerged which seemed particularly pertinent to my research are outlined below, indicating the related field drawn from in brackets in the related heading. At the end of this section, a brief outline of my phenomena is provided, showing how these informed my fieldwork, in terms of the elements or cues I was proposing to pay attention to in my observation and sensemaking (Weick: 1995).

2.7.1 The importance of tacit knowledge (Knowledge Management)

As in the IS field, two broad schools of thought can be discerned within KM (D'Eredita & Barreto: 2006), which seems to reflect a split along ontological lines. This crystallises in a focus by one school on treating knowledge as an asset or a resource, versus a focus on *knowing* as an activity or process by the other (Blackler: 1995, 1993). Both schools recognise the importance of tacit knowledge, although they conceptualise this very differently, with important implications for how they believe new knowledge may be created.

The widely cited resource focused school (Nonaka et al.: 2000, Nonaka: 1994) believe that creating new knowledge is fundamentally about the interaction between tacit and explicit knowledge. Nonaka (1994) identifies four patterns of such interaction: socialization, combination, externalization and internalization, positing a continuous 'spiral model' for creating knowledge, starting with individuals in an 'interaction community' or group (citing CoP as an example), then progressing to organisational and inter-organisational levels. Nonaka characterises knowledge creation as essentially about converting tacit knowledge, mainly to explicit knowledge, that can then be codified and shared as a resource. He distinguishes knowledge creation from learning, although his argument and distinction here are not clear, as he doesn't seem to exclude action-based or social learning. He may simply be pointing to a concern about more traditional forms of learning focused on acquiring existing codified or abstract knowledge but this is not explained or made clear.

Criticism of the widely cited resource view

Several researchers, from a socially constructed perspective, argue that Nonaka's notion of *externalization* and *conversion* from tacit to explicit knowledge is based on a misunderstanding of the nature of tacit knowledge (Tsoukas: 2005, Seely Brown & Duguid: 2000, Blackler: 1995). Tsoukas (2005) emphasises the complex nature of knowledge, and its implicit tacit human dimensions. Tsoukas (2005) criticises commonly circulated definitions such as Nonaka's for adopting a very narrow Cartesian view of knowledge and cognition, and not revealing a useful enough conception of its constituent components and how these interrelate. Instead, taking Polanyi as his starting point, he argues for an emphasis on the personal nature of knowledge, i.e. 'all knowing is personal knowing' (Polanyi quoted by Tsoukas & Vladimirou: 2001: p.974).

Based on a close reading of Polanyi's (1966) work, Tsoukas (2005) identifies the following essential elements of tacit knowledge:

- a coherent object of focus or phenomenon,
- comprising subsidiary elements, integrated subconsciously, and

- a person linking and integrating these components in pursuit of a purpose (realised in a focus for attention), using a semantic capacity and ontology to give meaning to the coherent whole.

He observes that tacit and explicit knowledge are intertwined and inseparable, therefore he first argues that it is impossible to *convert* tacit to explicit knowledge and, second, that any explicit knowledge will have associated tacit predicates that are inferred, based on experience, in the light of a relevant action context, purpose and values. As highlighted in the previous section, mounting evidence is being found within cognition research for these ideas of how tacit knowledge is created and proliferates (D'Eredita & Barreto: 2006).

In spite of Tsoukas' (2005) criticism, he and Nonaka nevertheless seem to agree on the importance of tacit knowledge and that the following aspects are important for knowledge creation:

- its action orientation or purpose,
- its situation within a specific context and 'interaction community' or community of practice,
- related reflection and sensemaking activities, and
- its social nature and the associated importance of dialogue, language and metaphor for collective learning, sensemaking and dissemination to occur.

A preferred emphasis on knowing

By contrast to the resource focused school, the social constructionist characterisation, as outlined by Blackler (1995), emphasises the process or activity of *knowing,* rather than abstracted knowledge as a resource, characterising *knowing* as:

- Mediated
- Situated
- Provisional
- Pragmatic, and
- Contested.

Tsoukas (2005) in turn stresses the 'ineffable' nature of tacit knowledge. He argues that the knower, focusing their attention on a focal target or purpose, is only peripherally aware of subsidiary particulars that may be relevant to their purpose or focal attention. Subsidiary particulars are assimilated through experience and practice and are interiorised over time, forming 'an unarticulated background' which influences and frames action but cannot be focused on *during* action. Instead, he argues that particulars can only be focused on during reflection on the activity, with a view to drawing attention to features of our action that may have escaped our attention during the action (which act as cues for interpretation and sensemaking). He therefore argues for the centrality of reflecting on practice and drawing attention to particulars or features of a phenomenon within a particular action context in order to generate new knowledge or insight. This seems consistent with Schön's (1991/1983) argument for the importance of practitioner reflection as a source of learning and insight.

Narrative knowledge

Given the time-bound, contextual, recursive and socialised nature of knowledge, Tsoukas (2005) argues for the importance of what he terms narrative knowledge, embedded in practice and constantly evolving through dialogue, reflection and practice, which he feels is likely to be neglected in institutional settings. He goes on to point out several paradoxes created by consistently privileging abstract, universal propositional knowledge and its related simplifying, rules-based approach to management. Instead, he sees both of these types of knowledge as relevant and existing on a continuum, where propositional knowledge and rules (grounded in tacit or implied predicates) are created to provide a consensus for action by providing a measure of certainty. He sees narrative knowledge as having the advantage of recognising the narrator, the context and its reflexivity, the narrator's and characters' motives or purposes, and the particular temporal context of the knowledge (i.e. not seeking universality). In doing so he stresses the critical role and use of language and dialogue, in order to facilitate making increasingly fine distinctions about a phenomenon, within a recognised action context. He

regards this as a defining characteristic of knowledge at individual and organisation levels, and argues for the importance of questions of epistemology at both levels.

The importance of dialogue

Tsoukas (2009) finds widespread support for the importance of social practices and social interaction for new knowledge to 'emerge', agreeing with Nonaka's idea of creating new knowledge through dialogue and the importance of using metaphoric language to facilitate this. Turning to research on dialogue and creative cognition, he theorises and richly illustrates how dialogue can give rise to new knowledge. In essence, he distinguishes *productive* dialogue (contrasted with *calculated*), describing it as collaborative exchanges to address mutually perceived 'strangeness' to generate new concepts or distinctions. When new distinctions are inter-subjectively accepted, these then represent new knowledge, which gradually gains wider acceptance and becomes part of what he calls 'the inherited background', forming an accepted knowledge context for future action and dialogue.

As part of this work on dialogue, Tsoukas (2009) points to the possible role and importance of what he terms *boundary artefacts* to facilitate productive conversations between actors or participants, by acting as 'an across-boundaries shareable framework, tool, object, or tangible demonstration' (p. 952). This seems a particularly useful concept for multidisciplinary teams (from different CoP) interacting to develop new insights. Tsoukas (2009) calls for more research on the dialogical creation of knowledge between different CoP.

2.7.2 The importance of framing and ambiguity (Sensemaking)

The sensemaking literature is cited by several of the social constructionist perspectives already outlined as influential in providing underpinning ideas and concepts for their work. However, Weick (1995) steers clear of providing a neat or simple definition of sensemaking, opting instead to provide a rich exposition of "the seven distinguishing characteristics that set sensemaking apart from

other explanatory processes such as understanding, interpretation, and attribution" (p. 17), with which it might otherwise easily be confused or equated. He explains sensemaking as a process that is:

- "Grounded in identity construction
- Retrospective
- Enactive of sensible environments
- Social
- Ongoing
- Focused on and by extracted cues
- Driven by plausibility rather than accuracy" (Weick: 1995: p.17)

Weick's sensemaking work contributes several key concepts and considerations in relation to how insights may emerge, in particular:

- The importance of *enactment* for meaning and the extraction of cues
- The distinction between uncertainty and ambiguity, and its implication that more data are only useful when addressing issues of *uncertainty* rather than *ambiguity*
- The idea of minimal sensible structures connecting cues with pre-existing frames in order to create meaning
- The impact of arousal on perceptions of context and its likely adverse impact on sensemaking (which may offer one explanation for the problem of information or data overload noted earlier in relation to big data).

Given the relevance of these ideas to my research in relation to a participant framing their inquiry, they are outlined in more detail in Appendix E.2.

Weick's work focuses largely at the level of the situated individual or group, essentially making sense of their context (most often organisational), attributing meaning to it in order to inform action. He makes an explicit connection to Lave and Wenger's (1991) work on situated learning and goes on to describe sensemaking's possible broader adoption as a *perspective*, as "a frame of mind about frames of mind that is best treated as a set of heuristics rather than as an algorithm" (Weick: 1995: p. xii).

The importance of sensemaking for IS

Given the pervasiveness of IT, Weick (1995) argues for the need for more interpretive research of IS in relation to sensemaking. He identifies several concerns in relation to IT and how these may impact on sensemaking. These centre on the limitations of rationalist, algorithmic IT approaches to anticipate all situations in a complex setting and their inability to facilitate reframing and identifying new, relevant cues.

As an important example of such work, he cites Orlikowski (1992), who draws on structuration theory to offer a socially constructed explanation of IT systems and how they are used. The ideas of institutionalisation and use she explores are consistent with Weick's (1995) and Wenger's (1998) characterisations of systems as reifications of practice. Subsequently, her work in this area has gone on to focus on issues of '*entanglement'* involved in tool and systems use and how these impact on framing and generating new knowledge (Orlikowski: 2007, 2006 & 2002, 2000).

The most important idea to emerge from Orlikowski (2007, 2006, 2002, 2000, 1992) and Weick (1995), in relation to deriving and discovering insight from data, is that systems institutionalise (or reify) the *designer's* thinking and assumptions at the time of developing the system, although this is subject to subsequent reinterpretation and repurposing by practitioners when using it. The extent to which these become fixed and inflexible are at the root of Weick's (1995) framing and sensemaking concerns.

Similar concerns may arise for data design and use, distinct from algorithmic, process or practice elements, in terms of framing the phenomenon the data purport to describe. For example, which elements or dimensions are relevant, may also be reified or institutionalised, thereby bounding the nature of the questions that can be asked of such data and what new knowledge it is possible to generate or discover.

2.7.3 The likely importance of research concepts

My own research as seeking insight from data

The research literature alerted me to the recursive nature of my research inquiry. I am searching for insight through the data I am collecting and analysing about initiatives that are also seeking to derive or discover insight from data. This highlighted rich possibilities for triangulation and reflexiveness, mirroring and contrasting my research with the phenomenon being studied (Davies: 2008). These aspects are addressed further in the following Methodology Chapter and more detailed Data Analysis Chapter.

My review of research philosophy and design literature also highlighted the importance of research *questions*, and that these evolve as one gains a clearer picture of the phenomenon of interest, migrating from *What* questions, via *Why* questions to *How* questions (Blaikie: 2007). Epistemological and ontological considerations also emerged as pertinent, helping clarify the exploratory rather than confirmatory nature and purpose of my intended research.

Absorptive Capacity and Path Dependency

At a slight distance from pure research thinking, the concept of *absorptive capacity* in relation to developing insights (Cohen & Levinthal: 1990) also seemed important. It resonated with ideas of path dependency and prior knowledge already highlighted by the knowing strand within IS (Shollo & Galliers: 2013). This idea is based on work on corporate research and development activities in the context of new product development, technology adoption, innovation and an organisation's related long-term competitive success. Cohen and Levinthal define the concept as follows:

> *"The ability to exploit external knowledge is thus a critical component of innovative capabilities. We argue that the ability to evaluate and utilize outside knowledge is largely a function of the level of prior related knowledge. At the most elemental level, this prior knowledge includes basic skills or even a shared language but may also include knowledge of the most recent scientific or technological developments in a given field. Thus,*

prior related knowledge confers an ability to recognize the value of new information, assimilate it, and apply it to commercial ends. These abilities collectively constitute what we call a firm's 'absorptive capacity' " (1990: p. 128)

Their concept clearly represents an *organisational* capability related to external knowledge and organisational learning. While their subsequent empirical research is quantitative in nature, their rich description of the concept draws on several ideas from cognition, associative learning from past experience and knowledge. They argue that *path dependency* provides a useful framework for understanding learning and new knowledge aquisition, and for the related importance of having a variety of expertise and experience available. This provides access to a wide variety of possible cues and to rich sets of possibly relevant prior knowledge. Such ideas seem consistent with subsequent social constructionist KM, sensemaking and learning ideas outlined earlier, in particular Vera and Crossan's (2003) proposition related to the role and importance of prior knowledge for learning.

Based on prior empirical work, Cohen and Levinthal (1990) argue that preconditions for successful learning and problem solving are the same, highlighting the following as important, which may also arise in a project context:

- Roles of 'gatekeepers' and/or 'boundary spanners' to bridge external and internal boundaries who through an outward focus, assimilate external knowledge, then translate and communicate it internally to relevant parts of the organisation.
- They characterise these touchpoints as 'interfaces' where the relative complexity of the knowledge/capability on the external side prompts the need to translate what is relevant for participants on the other. Indeed, they argue the distribution, prevalence and depth of expertise represents a major factor in the likely absorptive capacity over that interface, as well as whether or not specialised 'gatekeepers' or 'boundary spanners' are likely to be appropriate.

- A further factor related to this last question of specialisation is seen as the relative speed of change in the external environment and rate of knowledge development, where rapid change would argue against too much specialisation.
- They recognise the issue of considerable time lags being likely between assimilation of new knowledge, staff or technology, and then developing related internal organisational capabilities, subsequent use or related action.
- Competency 'traps' are also recognised as a potential barrier to absorptive capacity for new knowledge, where practices and approaches are successful and become entrenched or institutionalised.

Having explored these themes Cohen and Levinthal (1990) identify the following important characteristics of absorptive capacity:

- That it is cumulative and path-dependent. The organisation requires certain prior knowledge to be able to appreciate the significance of new knowledge and developments and to assimilate them. Otherwise external knowledge may simply be ignored and the organisation may become 'locked out' of a developing area, rapidly becoming unable to catch-up.
- Much of the relevant prior knowledge alluded to (and therefore absorptive capacity) is likely to be tacit and reside in a variety of individuals and specialists within an organisation. The authors recognise this will also make it difficult to measure, particularly in relation to how much would be appropriate to invest in it.
- Related to this idea of tacit knowledge, the absorption of 'know-who' knowledge is also cited as important (rather than just 'know-what' or content knowledge). This enables understanding the areas or units where this knowledge may be relevant and who in the organisation to inform
- Absorptive capacity influences expectation formation in terms of trends, their likely impact and appropriate possible organisational responses

Many of these ideas are reminiscent of Weick's (1995) ideas of enactment and selective attention to cues based on prior experience. Cohen and Levinthal's

exploration of communication between practice groups and the need for translation also connects with Wenger's (1998) ideas of economies of meaning evolving which require boundary spanning activities, practitioners and artefacts to bridge effectively.

The absorptive capacity construct clearly arises in a different context, focused on organisations as a unit of analysis, although Cohen and Levinthal accept an individual starting point for the concept. Nevertheless, it represents a useful contribution to thinking about how insight may emerge or be discovered.

2.7.4 The data project phenomenon in view

The literature and ideas outlined identified various aspects of pursuing insight that seemed relevant for my research. The accompanying description below, seeks to capture these to reflect a broad definition of what is meant by exploratory data initiatives or projects (used interchangeably within this thesis), which are the focus of my research. This depiction was created prior to undertaking fieldwork and was included in the research proposal document, prepared as a basis for discussion with potential research participant organisations.

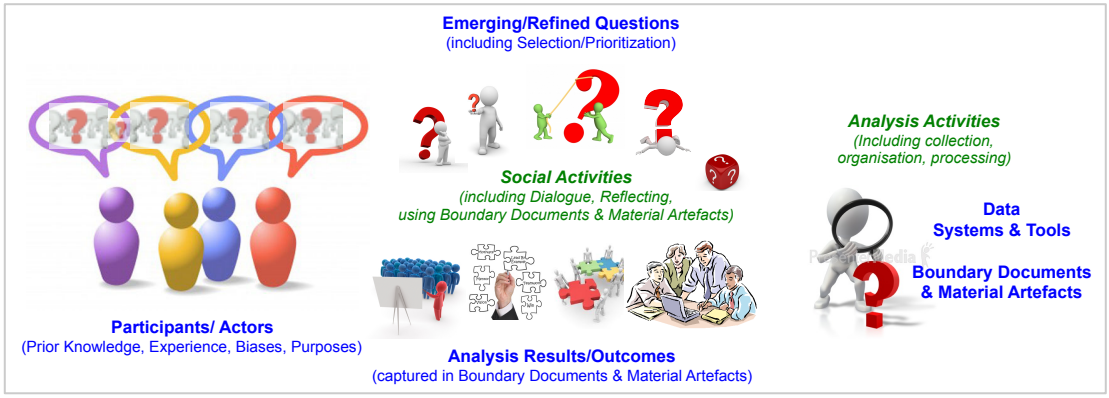


Figure 2-8 Phenomenon of interest within a Data Project Context

> *"We anticipate that the project team will engage in various activities (explicit and implicit) to come up with and gain consensus around the purpose(s) and question(s) they think are important and are likely to lead to valuable insights. They will also need to agree on what data is likely to*

be relevant and how this should be analysed. Existing theory and research suggests that participants' prior experience and knowledge is likely to be important factors in this process. A further dimension captured in the diagram is that the investigation or research activities themselves (e.g. data collection, classification or organisation, and processing/analysis), together with emerging preliminary findings and frameworks may also influence and help refine the questions and purposes pursued through learning." (Participant Research Proposal)

The *Green* (*italic*) elements highlight anticipated social processes or activities in focus while **Blue** (**bold**) elements indicate inputs, tools, artefacts and outcomes.

A central focus is on how exploratory projects are framed by participants, and the questions they pursue (implicitly or explicitly), recognising that they encapsulate effective situated learning objectives for the project. Framing and Questions also reflect the level of clarity and the prior knowledge about the phenomenon of interest. Blaikie (2007) posits the idea of a progression from posing 'what' questions, towards posing 'why' questions and ultimately addressing 'how' questions, reflecting increasing learning about a phenomenon of interest and increasing research complexity. Tsoukas (2005) argues that the ability to be able to draw increasingly fine distinctions about a phenomenon is a sign of increasing knowledge about the phenomenon, which should be discernable in the nature of the questions being posed about the phenomenon.

2.8 Literature that subsequently emerged as significant

This section explores two areas of IS literature that emerged as important during my data analysis, reflection and writing up my research. Both relate to my discussion and argument to disentangle data from the wider MSS artefact, particularly in the context of data initatives (in section 8.1.4).

2.8.1 Efforts to conceptualise Data, Information and Knowledge

As already highlighted in section 2.1, defining these concepts, core to the IS discipline, has proved problematic and contested for many years (Kettinger & Li: 2010, Checkland & Holwell: 1998). In the most recent effort to revisit this question noted, Kettinger and Li (2010) argue for the important role of (prior) Knowledge in deriving and discovering information from Data. However, Kettinger and Li (2010) attempt to conceptualise these phenomena and theorise relationships between them purely in explicitly normative and reductionist terms. They do not address or adequately reflect the human agency involved, nor the socially constructed nature of these phenomena. This is likely to result in fundamental problems for efforts to define them and theorise about them. Some of these problems are outlined in a critique below in relation to my research context of theorising how insight may be generated from data.

Missing human agency and sensemaking

Kettinger and Li's (2010) broad argument and definitions for key concepts reflect an information processing view of IS, systems and artefacts (Orlikowski & Iacono: 2001). In particular, they consider *Data* as "measurements or conditions of states about objects or events" (p.409), while *information* represents a "status of conditional readiness for [purposeful] action" (p.409). Their view of Knowledge seems particularly narrow, focusing on codified condition-action pairs; these don't acknowledge or address the important role of any necessary tacit precepts in relation to any codified Knowledge (Tsoukas: 2005). In their own terms, it also doesn't address boundary conditions that might apply to sets of such condition-action pairs.

They mention or acknowledge Data users, in terms of reinterpretation, and variations in interpretation of Data, through the application of different Knowledge. However, they don't address the role and processes of translation or authorship of (human) observation into Data terms using codes and applying prior Knowledge during this process. For example, Data *designers* or *capturers* are not addressed in their model or discussion, e.g. in terms of defining Data dimensions, categories, states, etc. in relation to a Phenomenon of interest.

Their model therefore omits Checkland and Holwell's (1998) idea of CAPTA or Weick's (1995) notion of paying attention to particular cues within a situation based on prior experience and Knowledge. On this latter view in particular, what is paid attention to and therefore captured is inescapably tied to prior experience and Knowledge recalled in the moment of observation and recording. To the extent automated Data capture is introduced (e.g. through sensors), the designer role becomes particularly important, as it can 'lock-in' design, reifying and introducing rigidity into what might be observed and captured about Phenomena. This increases the likelihood of unreflective assumptions being made, with the resulting risk of unintended consequences of Data reinterpretation and repurposing. This may have particular implications for sensor measurement and use given the trend towards an 'internet of things'.

Knowledge is similarly taken as given, 'factual' or objective, although allowing for it being provisional and subject to revision in the light of expert validation. This fails to recognise the complexity of various, often interacting, Knowledge domains, which is highlighted in the research findings and wider research about 'ontologies' (Ribes and Bowker: 2009). Such research about 'ontologies' focuses on efforts to make these explicit and relate them to each other but may suffer from similar challenges where it fails to acknowledge and address the inevitability of tacit precepts and limitations to making all Knowledge explicit (Tsoukas: 2005). Inherent category challenges highlighted by Star (2010) are also likely to arise with such an approach, reflecting Weick's (1995) argument about the importance of human sensemaking and reframing to address the impossibility of anticipating all eventualities in any codified logic.

Not acknowledging the related and implicit *design* process, is illustrative of a focus on *processing* (also implying automation) rather than the broader processual aspects of the Data Phenomenon, its design and use over time. This leads in some cases to putting forward problematic ideas, for example that Data have *inherent* meaning without any human sensemaker or observer involved. Similarly, they introduce the idea of 'processing *information* to produce higher-level *information'*, without defining how such levels are to be judged in the

absence of someone to make that judgement in relation to a particular purpose or objective. The ideas of Data quality and content richness touched on, are likely to be similarly contextual, given particular user specific objectives.

While the authors acknowledge the importance of user purpose and context, briefly mentioning structuration theory, they do not really engage with how this informs or impacts on their model. They also don't address how language fundamentally mediates meaning (e.g. semiotics), or engage with the social strands of the KM literature, such as Blackler (1995) or Tsoukas (2005). The latter may be particularly helpful in addressing a definition for Knowledge, and help clarify our thinking about Data within IS. Indeed, KM may consider Knowledge (or *knowing*) to be a core concept of '*their'* field and their 'territory' to define and theorise, rather than seeing it as a core concept of the IS field.

Interchangeable use of concepts points to confusion

Kettinger and Li (2010) highlight the interchangeable use of Information and Knowledge concepts by some researchers as problematic. However, they do not address the far more common use of Data and Information as interchangeable terms, both within the field and by practitioners. Both point to a lack of clarity or confusion about the underlying phenomena and the distinctions being drawn between them. By way of example, a recent call to theorise the IS artefact from a design perspective, argues for the recognition of an Information element for such artefacts. This argument equates information and Data in its related discussion (Lee et al.: 2015).

Let's first consider Kettinger and Li's (2010) approach to distinguishing these concepts. They ask the reader to consider different propositional statements as examples of what might be conveyed, to illustrate distinctions between them that they expect the reader will draw. Yet, this ignores a number of taken for granted, or implicit elements of prior Knowledge on the part of their readers. Some of this Knowledge relates to what they term *coding*, which can also be thought of as categorising or what Tsoukas (2005) refers to as making distinctions. Other areas of Knowledge drawn on relate to semantic and pragmatic meaning. For instance, in more straightforward, human terms, it

implies that the author and reader are both literate, share a common language and related usages and meaning. Even then, its implicit focus on the explicitly textual, doesn't allow for tacit elements that may be conveyed during an equivalent spoken conversation, e.g. emphasis or non-verbal cues, etc. Indeed, these may be difficult to capture in text, which highlights the role of the *medium* used for recording, communicating or transmitting Data.

For instance, while some non-verbal elements may be more available within a sound recording, and still more in a video recording, they may present trade-offs, in terms of relative storage (volume), as well as the speed of retrieval and interpretation, e.g. in the case of unusual accents. This illustrates how the choices of medium may reflect explicit or implicit choices about what to capture in relation to an observable phenomenon, i.e. Checkland and Holwell's (1998) idea of CAPTA. In turn, this influences what is available for subsequent engagement and analysis by a data user. Where an original observer of the phenomenon engages with such Data, it will also cue or trigger tacit memory of the original observation or experience (e.g. their emotional response at the time, etc.), not available to others. This may also offer original observers opportunities for validation and triangulation during interpretation. Indeed, this was my experience when reviewing and analysing my research transcripts and recordings, described in later chapters.

What purpose might these distinctions serve?

A primary consideration is the benefit of being able to make distinctions about and between these concepts. This should, in turn, be reflected in how these distinctions are used or employed by researchers or practitioners. That will have a significant bearing on which dimensions (or layers and related theory) are relevant. For example, are we interested in interpretation and action or in technical data transmission? For a model or theory in this area to have broad explanatory power, as well as addressing parsimony, it needs to satisfactorily address these concepts in both automated and human settings. Ideally, concepts and distinctions should be broadly useful across different technical

and human contexts, explaining both, or should at least make its scope and boundary clear.

By way of grounding this critique, Figure 2-9 presents a rich human context envisaged, which any concepts and theorising would need to address. This depicts a human observer of a phenomenon within the observable world, who may also choose to capture or engage with related data, or communicate with others about a phenomenon, all the while paying attention to particular cues, filtering and relating what is observed in terms of embodied prior Knowledge and experience of what might be relevant to the person's purpose in the moment. For simplification purposes, this is drawn from the perspective of an individual rather than the group level phenomenon addressed by my research. This depiction represents a modification of Checkland and Holwell's (1998) context for IS.

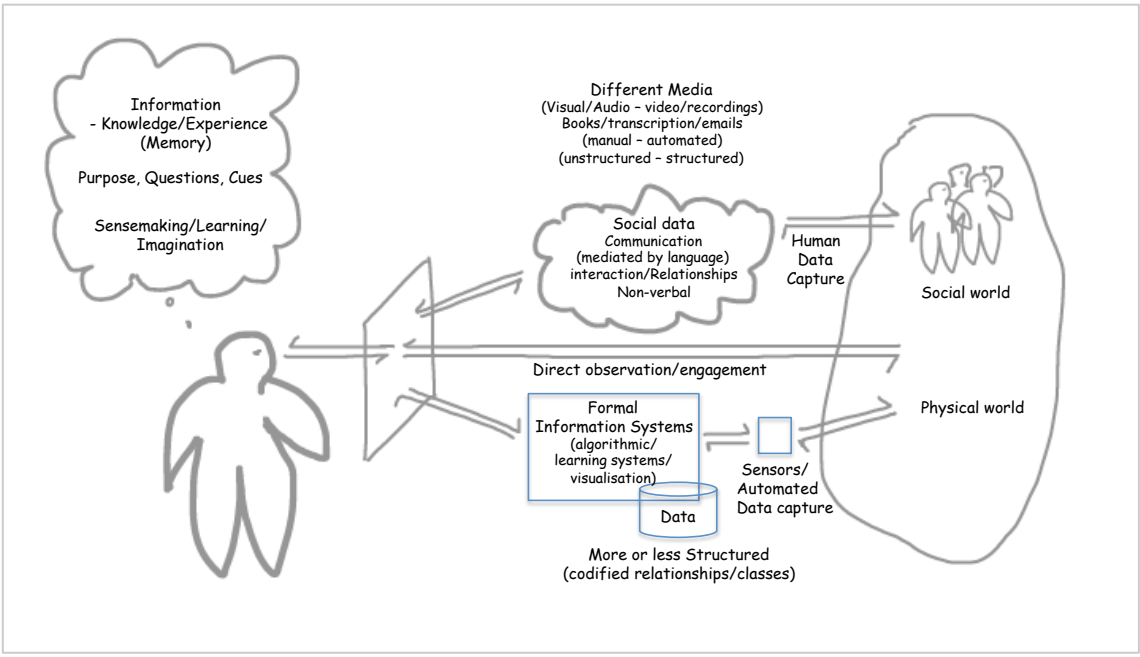


Figure 2-9 Human context for theorising Data, Information and Knowledge
(Modification of Checkland & Holwell: 1998: p.98)

Considering the above in relation to the concepts being defined, theorised and distinguished, it highlights how data are something fundamentally different from information and Knowledge. If we think of what is clearly discernable about

these phenomena to an observer of this situation, data clearly exist as artefacts in some tangible form, even if digital. Data are independent of any phenomenon or human observer.

On the other hand, we cannot clearly discern information or Knowledge, which are ineffable and subjective, embodied within the human observer (and fellow observers). Basing the distinctions between these concepts on their inherent nature and where they reside, makes distinguishing them far easier and clearer. It also points to a natural reason for the difficulty in discerning information and Knowledge. Checkland and Holwell (1998) implicitly recognise this in their indicating more complex, longer-lived structures being evident in Knowledge as differentiators from information. This, in turn, points to these concepts occurring on a continuum rather than being clearly distinct phenomena.

This seems to provide a far more useful starting point for theorising these concepts, than adopting a more abstract starting point and distinction. It points to the very different factors and considerations at play when seeking to investigate, manage or influence these phenomena. Indeed, this offers an explanation for the far richer distinctions we can and do draw about data within IS than the more ineffable and embodied information and Knowledge.

This emphasis and bias towards the more tangible also extends to our focus on codified forms of Knowledge, both within IS and by mainstream theorists and researchers in KM, reflected in a dominant resource based view of Knowledge. Based on the scenario outlined, such codified forms of Knowledge might more correctly be thought of as more complex and relatively unstructured forms of data. This would certainly provide a more consistent and straightforward starting point for theorising, distinguishing and managing these phenomena.

The above example and discussion illustrates the interrelationship and potential relevance of various areas of literature, such as cognition, communication and textual interpretation when seeking to theorise these concepts, and to make distinctions about them or between them. It also points to the value in the IS discipline working closely with social constructionist researchers within KM and situated learning. For example, drawing on insights from the social

constructionist strand of KM, it may be worth focusing on use, or a processual notion of *informing*, rather than persisting with a resource based view of these concepts. We may even be able to simply adopt their notion of *knowing*, if we take seriously the idea that Information and Knowledge occur on a continuum of complexity and longevity. Indeed, some IS researchers have already adopted knowing as a useful approach (Newell & Galliers: 2006).

2.8.2 Sociomaterial efforts to theorise the IT Artefact

This section examines an emerging and important strand of IS research which seeks to recognise the interrelationship between the social and technical or material aspects of IT artefacts. My research did not specifically seek to theorise IT artefacts, nor did it adopt a sociomaterial starting point or approach. Nevertheless, it has emerged as relevant for my argument to distinguish and theorise data as a distinct IT artefact, separate from a wider system artefact that it may be embedded within. A brief synopsis and evaluation is provided of key central debates noted that emerged as relevant to this argument.

An overview of the sociomaterial journey so far

Orlikowski (1992) has long championed the need to better theorise the IT artefact holistically within a situated, social context. Initially, she adopted a structuration lens with which to examine IT systems. On this basis, she argued for recognising a duality of reified structural elements and tacit reinterpretation during use of systems over time.

In later research, Orlikowski and Iocona (2001) provide an analysis of different conceptualisations of the IT artefact used within IS research and literature, identifying the following broad views or categories: a tool view, a proxy view, an ensemble view, and a computational view. They also recognise the absence of the IT artefact in a significant proportion of studies they examine (during the 1990s), which they term a nominal view. They are concerned that the IT artefact may “disappear from view, be taken for granted, or is considered unproblematic once it is built and installed” (p.121). Based on this, they argue for more focus on the core IT artefact within IT research, and theorising it explicitly.

This has led to a rich strand of sociomaterial research within IS (Cecez-Kecmanovic et al.: 2014), which seems to be gaining momentum. Recognising three broad theoretical roots for sociomaterial thinking: sociotechnical systems, actor network theory, and practice theory, Cecez-Kecmanovic et al. (2014) identify two main ontological stances or poles within sociomaterial research, which they term relational and substantialist. The relational stance, associated with Orlikowski and Scott (2008), which goes beyond Orlikowski's (1992) initial dualist argument, emphasises the inseparability of the social and material. It also emphasises the performative nature of practice and interactions between human and material agents, considering both as a nexus of relationships. On the other hand, the substantialist stance stresses that the human and material are independent, although influencing each other. Cecez-Kecmanovic et al. (2014) identify several variants of these two 'polar' perspectives, as well as the notion of strong and weak relational approaches.

Relevance for my research and argument

While my research did not adopt a sociomaterial starting point or approach, it has much in common with sociomaterial research. This is evident in its interest in the interaction between the technical and social, and their interdependence. It is also evident in my focus on the IT artefacts at work within particular practice contexts, and an emphasis on theorising these artefacts and their impact within the contexts studied.

One of the essential debates identified between variants of sociomaterial thinking relates to the degree of agency that can and should be attributable to material and human actors or actants, and the need to distinguish these different actants. For example, human actors having intentional agency is seen by some as an important distinction between human and material actants.

As already highlighted earlier, sociomaterial research to date has not specifically focused on theorising data IT artefacts or on data analytics related initiatives or contexts. My research examines the use of different tools and data by different practitioner groups, and incidentally sheds some light on different kinds of agency at work and how this may shift over time within particular

contexts. This discussion touches on differences noted between human sensemaking and data processing based on case observations. It also distinguishes data and tool elements of IT artefacts (often integrated within an MSS or IT artefact) and its implications.

Particularly relevant to my discussion, is the argument put forward by Leonardi (2013), emphasising the ability for the material to persist over space and time, independent of humans,.This is revisited and explored in section 8.1.4 of my Discussion chapter in relation to the collection, use and evolution of data over time noted within the cases examined.

Finally, Cecez-Kecmanovic et al. (2014) recognise the importance of making sociomaterial research relevant to practitioners, and the danger of resulting theory being inaccessible to them. In this regard, one concern about some strands of sociomaterial research is their emphasis on explanatory research, addressing only how questions rather than also seeking to address why questions, which may have greater normative or practical value. This explanatory focus is acknowledged when actor network theory approaches are adopted.

My discussion will seek to address this by engaging with the existing normative view of technology within the causal MSS model, from a more social constructionist or sociomaterial perspective. As these may be more accessible, familiar and useful to practitioners, in their aim to provide normative guidance rather than only seeking rich, explanations, this may generate greater research impact on practitioners.

3 METHODOLOGY

Given the relative lack of prior research into the processes and activities at work in deriving or discovering insight from data in data initiatives, this research is essentially exploratory in nature, seeking to synthesise, extend and refine theory. Therefore, an abductive research strategy was adopted (Blaikie: 2010).

The rationale for adopting this strategy is outlined in the following sections, highlighting the inherently human and social nature of the phenomenon, as well as the research focus on studying it within an exploratory project context, prompting an ethnographic approach. This approach is argued to encompass a wide range of active, immersive participation by the researcher (Baskerville & Myers: 2015, Blaikie: 2010, Eden & Huxham: 2002), facilitating rich access to the phenomenon. The approach is also recognised as a particularly in-depth research method, well-suited to research focused on human, social and organisational aspects of IS (Baskerville & Myers: 2015).

An idealist ontology (Blaikie: 2010) informs my research and the ethnographic approach adopted reflects an engaged researcher stance, most closely aligning to what is termed ‘Mediator of Languages’ (Blaikie: 2010). In ethnographic terms, the research focus is therefore etic (an external perspective) rather than emic (an insider perspective) (Hammersley & Atkinson: 2007). This reflects the pursuit of theoretical meaning and explanation and the sensitising role of theory.

A summary overview of the research design adopted is provided in Appendix A. The chapter starts by providing a more detailed rationale for adopting the approach outlined above, in preference to other possible approaches. The rest of the chapter goes on to outline my data sources, case selection, data collection and briefly touches on my data analysis approach, as well as the role of theory and reflexiveness.

3.1 Rationale for Research Design

This section starts by providing more detail on my research perspective and stance adopted in the field. It then goes on to present an argument for the

research strategy and approach adopted. This is based on the nature of the research question posed, and an evaluation of various methodological options.

3.1.1 Research Perspective and Stance adopted

My perspective reflects a personal commitment and the nature of the research being undertaken. My personal ontological starting point is that the nature of social reality is fundamentally socially constructed. This position is clearly near the Idealist/Nominalist end of the spectrums outlined above by Blaikie (2007) and Easterby-Smith, Thorpe and Jackson (2008) respectively. I identify with Chia's (2002) argument for a more processual view of social reality and ideas of structuration in social theory, based on Giddens' (1984) ideas.

The nature of the phenomenon and my research is essentially concerned with how meaning is established by participants (i.e. socially constructed) and consensus achieved (or not) as to purpose and questions pursued. It also focuses on what and how they learn from the pursuit of insight, both about a phenomenon of interest and how to go about this activity or practice. Clearly the nature of the phenomenon in view points to a social constructionist or idealist starting point for my research. The processual or longitudinal nature of the research, examining what is learned over the course of a project or initiative, as well as its focus on socially situated participants, pointed towards related research paradigms of Ethnomethodology, Action Research or Cooperative Inquiry.

A socially constructionist epistemology also aligned well with the key elements of research and authors that I identified in my review of relevant literature as important starting points to build on (as per Chapter 2). In particular, this was supported by Easterby-Smith and Lyles' (2003) identification of a particular strand of such thinking and research within Learning and Knowledge Management, which they characterise as social constructionist, in terms of sharing a set of assumptions.

Researcher Stance

Blaikie (2010: 50) identifies several possible stances that a researcher can adopt towards the research and participants. These are summarised in Table 3-1, together with key authors they are associated with or originate from.

Blaikie argues that these stances are often associated with particular research strategies, which I considered when selecting an appropriate research strategy and method(s). Based on the descriptions provided, the stance that most closely resembled mine, was that of Mediator of Languages, although I recognised that this stance may overlap with that of Reflective Partner and Dialogic Facilitator.

Table 3-1 Researcher Stances (based on characterisations by Blaikie: 2010: 50-54)

Stance	Brief Description	Key Authors
Detached observer	Represents the traditional '*scientific*' stance, grounded in the belief that reliable knowledge can only be produced based on *objective* approaches which address the risk of researcher values and preferences influencing the research, particularly during data collection. This approach is widely criticized.	-
Empathetic observer	Also tries to achieve the same objectivity but stresses that the researcher needs to grasp social actors' subjective meanings in order to understand their actions, requiring the researcher to put themselves in the same position as participants. This stance is commonly referred to as *verstehen* (*i.e. understanding*).	Weber (1964) Outhwaite (1975)
Faithful reporter	Evolved from the second but is much less detached, including possible immersion. The researcher's aim is to facilitate or allow participants to 'speak for themselves' and faithfully present their point(s) of view. This is often referred to '*naturalism*' and requires study to be faithful to the phenomenon, conducted in its natural state, and its description to be recognizable to participants for validity (often achieved through participant validation).	Lofland (1967) Blumer (1969) Matza (1969) Denzin (1971) Douglas (1971) Guba (1978)
Mediator of languages	An extension of the third position but which rejects the notion of detachment. The researcher is seen as interpreting participant accounts and *translating* between everyday or lay language into scientific language, thereby unavoidably introducing an authorial voice, making objectivity impossible. Social, geographical and historical locations, as well as researcher interests and assumptions are recognized as important impacts on interpretation.	Giddens (1976) Gadamer (1989) Geertz (1988)
Reflective partner	Associated with *critical theory*. The researcher is committed to the *emancipation* of participants from whatever oppression experienced. Rejects the 'objectivist illusion' and believes social and cultural reality is pre-interpreted by participants through a dialogically constructed shared framework of meaning (which changes over time). The researcher is seen as a co-participant. Another version is associated with feminism, involving conscious partiality and requiring conscientizing participants and the researcher.	Habermas (1970) Mies (1983) Freire (1970)
Dialogic facilitator	Seen as a *post-modern* culmination of the last two stances, the researcher is regarded as another participant, rather than an expert, taking elements of the last two stances. The researcher seeks to reduce their influence by facilitating a 'variety of voices' to emerge. The emphasis is on a *dialogue* between the researcher and researched.	Fontana (1994)

The key distinction for me between these stances, and the reason for my identification with the former, Mediator of Languages, was that I was seeking to present a technical or theoretical explanatory account of what was happening with a view to theory synthesis, refinement and extension. I was not seeking to use or adopt critical theory. Importantly, all of the stances considered reject the idea that it is possible for a researcher to be detached, prompting the need for Reflexiveness on the part of the researcher.

Research Strategy and Methodology

Blaikie (2010) recommends that research strategies be adopted on the basis of the researcher's perspective as well as the nature of the research questions being posed. He argues that questions can all be broadly grouped into three types:

- **What** questions – these require a descriptive answer about a phenomenon in terms of its characteristics, patterns.
- **Why** questions – which seek either causes of, or reasons for, the existence of certain characteristics or regularities noted in connection with the phenomenon.
- **How** questions – concerned with how to bring about change and produce certain outcomes through interventions.

These are linked to different research purposes and in turn to different research strategies or logics of inquiry, depending on your perspective. Blaikie (2010: pp.69-70) identifies three broad categories of research purpose: description, explanation/understanding and change. He goes on to explain the connection between research questions, related purposes, and four broad research strategies (or logics of inquiry), as summarised in Table 3-2.

Table 3-2 Research Strategies, Questions and Purposes (Blaikie: 2010: p.105)

Research Purpose	Research Strategy				Type of Research Question
	Inductive	Deductive	Retroductive	Abductive	
Exploration	***			***	What
Description	***			***	What
Explanation	*	***	***		Why
Prediction	**	***			What
Understanding				***	Why
Change		*	**	**	How
Evaluation	**	**	**	**	What and Why
Assess Impacts	**	**	**	**	What and Why

Key: ****= major activity, ** = moderate activity, * = minor activity (these are indicative only)*

While my research question is written as a ***How*** question, my purpose is to explore and provide a description of a phenomenon that is not well understood, nor widely researched in my proposed context, with the intention of achieving a better understanding (i.e. Basic Research), pointing to what Blaikie would describe as a ***What*** question (e.g. in what ways are questions used, to what extent, etc.) and ***Why*** this might be, in terms of gaining a better understanding of underlying reasons. Therefore, considering and analysing my research question and associated purposes, adopting an Abductive Research Strategy (Blaikie: 2010) or a Qualitative Constructionist strategy (per Easterby-Smith et al.: 2008) was considered appropriate to answer the research question posed, and was consistent with my earlier choice on the basis of my broad ontological and epistemological starting assumptions.

Methodological Alternatives Considered

I found Easterby-Smith et al.'s (2008) summary of methodology elements associated with particular epistemologies a helpful starting point, clearly identifying with those associated with Constructionism, as illustrated in Table 3-3.

Table 3-3 Methodological implications of different epistemologies in social sciences (Easterby-Smith et al.: 2008: 63)

Elements of methodologies	Social Science Epistemologies		
	Positivism	Relativism	Constructionism
Aims	Discovery	Exposure	Invention
Starting points	Hypothesis	Propositions	Meaning
Designs	Experiment	Triangulation	Reflexivity
Techniques	Measurement	Survey	Conversation
Analysis/Interpretation	Verification/falsification	Probability	Sense-making
Outcomes	Causality	Correlation	Understanding

However, Easterby-Smith et al. (2008) identify a wide range of possible qualitative ***methods*** that can be used within the broad Contructionist Methodology, leaving considerable room for choice as to specific methods to adopt. Here, I found Gummesson (2000) more useful to clarify my thinking. He emphasises the importance of adopting methods that provide the best *access* to the phenomenon of interest. In my case the phenomenon relates to situated individuals and their practice of pursuing insight from data within an exploratory data initiative or project team setting.

Gummesson (2000) focuses on researchers and consultants in management research, identifying their similarities and differences and the value of reflecting on personal practice. He uses the metaphor of an iceberg (see Figure 3-1 on the next page) to argue for the central importance of access to practice and the idea of a hierarchy of increasingly good access.

Gummesson (2000) explains his notion of different research approaches providing different levels of access as follows:

- Traditional positivist approaches and methods are seen as fragmented and prone to misunderstanding, given the limited access to the phenomenon of interest, equated with a helicopter observation of an iceberg.

- Traditional qualitative informal interviews and observation provide much richer access to the phenomenon, equated with at least climbing onto the iceberg above the water line.
- Participant observation, or what he terms action research/action science, provides the richest source of insights, equated with much closer inspection, including below water-level.

Figure 3-1 The Iceberg Metaphor for Access

(Based on Figure 2.3 - Gummesson: 2000: p.36)

Gummesson (2000) makes a strong argument for action research achieving the richest access to management thinking and behaviour. The soft systems thinking strand of research in IS also uses a form of action research (Checkland: 1999, Checkland & Holwell: 1998) and was pivotal in my early thinking about how to undertake this research.

Ethnographic research (another 'full immersion', constructionist approach) is also widely represented in the work of researchers whose work influenced my research (Stigliani & Ravasi: 2012, Tsoukas: 2009, 2005, Orlikowski: 2002, 1992, Wenger: 1998). Ethnography increasingly encompasses researcher participation, as well as a more traditional observational stance (Baskerville &

Myers: 2015, Blaikie: 2010), although requiring the researcher to be reflexive about their presence and involvement (Davies: 2008).

Based on the research methods literature reviewed and described above, the following methods emerged as the most promising. They all facilitate good access to my phenomenon of interest and the ability to study its longitudinal, processual aspects:

- Case Study Research
- Action Research
- Ethnography

These overlap, especially in the nature of data collected, how data can be analysed and their criteria for quality and validity. Case Study Research doesn't really allow for the same level of researcher 'immersion' compared to the other two options (Harrison: 2002). While it can include an element of participant observation, this is not typical and certainly not the main emphasis or focus, so this approach was discounted. Action Research seemed promising but typically requires an *intervention* by the researcher, or their participation, seeking to *change* the phenomenon of interest (Blaikie: 2010, Eden & Huxham: 2002). Given that I had no plan to make discrete interventions and my research was more exploratory, descriptive and explanatory in nature, an Action Research approach didn't seem a perfect fit.

On the other hand, an ethnographic approach facilitates immersion and also encompasses a wide range of participation (Baskerville & Myers: 2015), i.e. offering more flexibilitiy. Baskerville & Myers (2015) argue that it is well suited to IS research focused on human and social aspects of a phenomenon, situating a phenomenon in its social and cultural context, and offering a particularly in-depth method.

In terms of their description of a range of ethnographic approaches that could be adopted, the approach adopted here represents what they term a more traditional 'anthroplogical' approach, rather than a design ethnographic approach, even though some active participation was envisaged. While the

anticipated duration of my fieldwork was shorter than more traditional ethnographic studies, this is balanced by its scope also being narrower than typical cultural ethnographies. Therefore, on balance, my clear focus and emphasis on longitudinal participant observation over the course of a project or initiative, pointed to taking an Ethnographic approach as being the best methodological fit.

3.2 Data sources and Case Selection

In view of the socially constructed nature of the phenomenon, I considered it important to gain an in-depth understanding of participant perspectives, their sensemaking and interaction over time in the context of a particular project, ideally in a natural setting. Hence, three main sources of qualitative data were envisaged:

- Participants (including key stakeholders) (anonymised throughout)
- The Researcher (as theoretically sensitised Observer and Participant)
- Project Documentation and Artefacts (e.g. minutes, requirements documentation, business cases, interim and final reports and presentations)

Rich organisational contexts were sought, to elicit multidisciplinary project team interaction dimensions anticipated across different departments and CoP within such initiatives. Projects were selected which explored new, unfamiliar sources of data rather than more straightforward projects where the goals, data and phenomena where all well-understood, i.e. the research focused on what could be termed exploratory rather than exploitative initiatives (March: 1991). Two such organisations (both based in the UK) were found to represent interesting and contrasting contexts: GoCouncil and InfraDig. These cases are introduced in Chapter 4.

In searching for contrasting cases, relative size and complexity was one feature considered important, with InfraDig far larger than GoCouncil and more complex in terms of organisational relationships involved. Another was to look at contrasting settings in terms of domain knowledge and the practitioner focus or

objectives for the data initiative. In GoCouncil's case this mainly related to market and customer research, broadly within the Marketing domain and working with Marketing practitioners, while InfraDig was firmly in the engineering domain, working with engineering practitioners. In the event, the two cases also contrasted in terms of the overall timespan for the initiatives, with InfraDig's initiative spanning a much longer timeframe (several years) than that envisaged at GoCouncil (an annual budgetary timespan), especially in relation to realising insights, which was in effect deferred beyond my field research involvement.

Such contrasting settings and initiatives were thought to offer rich possibilities for triangulation during data analysis and reflection, and indeed have done. This improved the likelihood of generating findings and insights that may be useful across different data initiatives, while stopping short of aiming for generalizability.

3.3 Data Collection

Data were collected in a natural setting, with myself immersed on site in the two case-organisational settings for a period of 6 months (between January and July 2013). I typically spent one day a week at the GoCouncil site and two days a week at the InfraDig site. Three sources of qualitative data were used: project participant interviews, researcher observation and participation, and project documentation or artefacts.

3.3.1 Orienting semi-structured participant interviews

In depth, semi-structured interviews were held with several project participants in each case to gain a participant perspective on their 'framing' of the data project, their roles and the extent to which this might reflect their experience, study or practice backgrounds. The interviews also sought to identify participant views of the questions being addressed by the project, or underlying learning objectives where questions were not explicit.

Based on these objectives and the preliminary view of the phenomenon outlined in section 2.7.4, the following specific areas were covered with the participants:

- Their understanding of the purpose(s) of the initiative
- The question(s) they were seeking to answer
- What data and tools they envisaged being used (and why)
- Their prior knowledge and experience, if and how this would be used during the project
- Their role within the initiative or as part of the team

A related interview protocol was developed ahead of the fieldwork, as part of formulating the research design (see Appendix B). When conducting initial interviews it was found that asking participants about potential challenges they anticipated for the project was a good way to get them to engage, reflect and open up in the interview about their project perspective and framing, so this was added.

3.3.2 Participative involvement

My involvement in both cases was participative, contributing ideas, often working jointly with other participants to develop project artefacts, reviewing and feeding back on documents and helping plan, run and facilitate workshops. For the duration of my project involvement, participants were observed during project meetings, informally between meetings and working together with some participants on tasks.

My involvement encompassed some *in situ* data gathering and related analysis, which was shared with participants, and is described in more detail in the next two chapters, which set out my case involvement and my approach to data analysis.

3.3.3 Data and artefacts collected

Notes were taken during interviews and meetings attended and they were recorded (wherever possible and practical), after gaining consent from participants. These were transcribed or summarised to facilitate subsequent

review. Formal meeting outputs, relevant project documentation and other artefacts were also collected (where feasible), e.g. requirement specification drafts, project update reports. Data were uploaded into NVivo v.10 (simply referred to as NVivo in the rest of the thesis) for tracking and the initial phase of data analysis.

Of particular importance were researcher field notes, captured in a research diary (Singh and Dickson: 2002, Emerson, Pretz & Shaw: 2001). Field notes were subsequently recaptured into electronic Microsoft Word form, and then uploaded into NVivo to facilitate further reflection and coding. It was not possible to record the facilitated workshops given the number of participants, challenges in obtaining permission and technical feasibility, although copies of workshop outputs and some photos of post-it wall-charts, were retained for reference and analysis.

Across both cases, 48 days were spent on site undertaking direct observation (with a corresponding number of field note journal entries). While on site, 14 in-depth participant interviews were conducted, 34 project meetings were attended, 23 additional meetings were held or attended, 3 workshops were (co-) facilitated, 49 artefacts collected and 19 project artefacts or outputs were (co-)produced. Of the meetings attended, 40 out of 71 (56%) were recorded.

A more detailed summary, broken down by case, is provided in Appendix B, together with supporting lists of interviews, meetings and artefacts.

3.4 Data Analysis

My data analysis seeks to provide a firm and transparent grounding for the explanatory accounts of the phenomenon (Van Maanen: 2011, Hammersley & Atkinson: 2007). This was based on a synthesis of the data collected rather than an emphasis on 'pattern-seeking' analysis and tabulation (Singh & Dickson: 2002, Miles & Huberman: 1994).

With this in mind, various strands of analysis and methods were adopted, primarily to facilitate and aid reflection through iterative engagement with the data, from various starting points and at different levels of analysis to illuminate

different aspects of the phenomenon, and create opportunities for triangulation (Singh & Dickson: 2002, Eden & Huxham: 2002). These approaches are explained in more detail in Chapter 5, after first introducing and providing some background to the two cases in Chapter 4.

3.5 Reflexiveness and the role of theory

As outlined earlier, theory played both a sensitising and a creative, dialectical role (Blaikie: 2010) during data collection and analysis. The sensitising role was most explicit in using the CoP framework (Wenger: 1998) and sensemaking (Weick: 1995) as a priori coding structures during data analysis. The dialectical role is most evident in the areas of literature which emerged as important to my argument to disentangle data from the wider MSS artefact, as outlined in section 2.8. In the interests of clearer recoverability (Checkland: 1999), these areas of literature were kept distinct from those that informed and helped frame my research inquiry and entry to the field in my literature review section (chapter 2).

Reflexiveness was also recognised as important and as occurring at various levels, prompting a pervasive approach. I sought to document it throughout my research as follows:

- maintaining a research journal of reflections throughout my Research
- capturing field notes into NVivo to make them available for analysis
- capturing reflections during transcription, summarising and coding, using a combination of memo notes in NVivo, addendums to meeting summaries, and in a data analysis notebook

These resources also represent data in their own right about my sensemaking of the phenomenon observed, generating insights from the associated data collected. This has offered a useful point of triangulation and another avenue for coding and reflection: about my role as researcher, the phenomenon as I understand it over time, as well as my own process of reflection, coding and deriving and discovering insights from my data. As such, it was used dialectically (Davies: 2008) and creatively to generate insights, seeking

opportunities for ‘triangulation’ and employing Mirroring and Contrasting approaches, in particular during such reflection (Singh and Dickson: 2002).

Reflecting arguments by Van Maanen (2011) and Tsoukas (2005) to acknowledge and reflect the narrator (and observer) during writing, I have sought to reflect and make my reflexiveness transparent by adopting a personal, first person narrative style throughout the thesis. My reflexiveness is particularly prominent in the confessional accounts of both cases in Chapter 6. In Chapter 7 too, I have sought to include reflexive elements during the cross-case comparison, treating my own research as a further case for comparison, highlighting these reflexive elements by placing them in *italics*.

4 CASES

This chapter introduces the two contrasting cases selected, providing some background context for each. A fuller explanatory account, or thick description, of what was observed in each case, is provided in Chapter 6.

4.1 Introducing GoCouncil

4.1.1 Council Background

GoCouncil manages a ‘district’ of 350 square miles, with a population of approximately 180,000 people. It is in a wealthy rural setting in England, within commuting distance of two large urban centres and is rated as having a good quality of life. As a district council it is part of a three-tier local government structure, as part of a larger ‘county’, managed by a county council, and sitting above two or three local municipalities, for large towns, and several parish councils, which manage the affairs of small villages.

The district council has a small executive team. Policy decisions are made by a board of elected councillors, advised by the executive team. It lists the following broad service areas on their website:

- Environmental Health and Licensing
- Planning and Building Control
- Housing
- Leisure
- Business services
- Benefits
- Parking
- Household and Commercial recycling and waste
- Community Safety (distinct from policing, which is separate)

GoCouncil, like most branches of UK local government, is receiving a steadily and significantly reducing central government grant towards covering costs. They have been relatively successful over the last few years at finding savings, to make up the shortfall without raising local council taxes, but the prospect of

significant continuing grant reductions is a critical concern for the council CEO. The easy wins, in terms of savings, have already been secured and there is no political appetite to increase local council taxes among the elected council board members. In fact, council members have grown used to the council consistently meeting progressive grant reductions, creating an expectation that they will be able to continue to do so indefinitely.

4.1.2 Business Model and Market Insight initiatives

Given the scale of the challenge, Alexander (the CEO) has long ago realised nothing short of a transformation of GoCouncil is required. Therefore, a year ago he instituted a programme to do this, headed up by Albert, one of his planning managers. Albert had recently completed a public sector MBA programme and was looking for a new challenge. While Albert doesn't have a project management background, he knows the council and personalities involved. He also seems commercially minded, which is rare at GoCouncil. Besides, the cost of involving an experienced external project manager or consultants is prohibitive for a council of their modest size, even though it serves a relatively wealthy rural area.

With this in mind, the CEO has instigated a programme to generate and implement business model thinking to revamp existing services or introduce new services. He hopes business model thinking will encourage entrepreneurial behaviour, engage teams in a greater commercial awareness of the costs of delivering particular services and spark ideas for generating income and smarter, more cost-effective ways to deliver services. Some units, like Leisure, with particularly large shortfalls and grant funding reliance need to shift their traditional approach of spending an allocated budget optimally, towards a more entrepreneurial approach to find other sources of funding (e.g. partnerships), generating revenue from their activities to achieve their social goals and outcomes. While this programme had been running for one year as I began the study, little progress had been made with almost no new business model proposals presented to the board, let alone implemented and securing new

revenue streams or cost savings. The funding gap remains largely unmet, putting increasing pressure on the executive team to show results.

Patricia and her Marketing team have positioned themselves to help by providing vital customer and market insight to various service teams as they formulate new or revised business model proposals. She is concerned at the quality of ideas emerging from initial discussions and draft proposals she has seen. They show little or no awareness of the related customer target market, likely competitor offerings, dynamics and price points. With this in mind she has got Board support for a project to provide market insight advice to the departments and functions. She has appointed Andrea, a relatively new member of her team, with experience from another council that has also sought to make savings, to lead this market insight effort.

She is certain that Acorn household data (a consumer classification of UK household demographic census data, http://acorn.caci.co.uk) represents a useful resource that can provide insights to the various teams in relation to their target market for various existing and new services. Her underlying assumption is that the council should try and identify and target households with more disposable income (e.g. the Flourishing Family Acorn category). Acorn data are also a cheap, existing resource that is easy to leverage,.

The initiative is summarised pictorially in Figure 4-1, highlighting the facilitation role played by the Marketing team in relation to others, positioning the use of the business model framework and Acorn household data, to help prioritise, refine and inform service proposals for different functions that could be put forward to the board for approval. The workshops also typically identified further market insight questions and requirements that could further strengthen proposals by grounding and testing key assumptions.

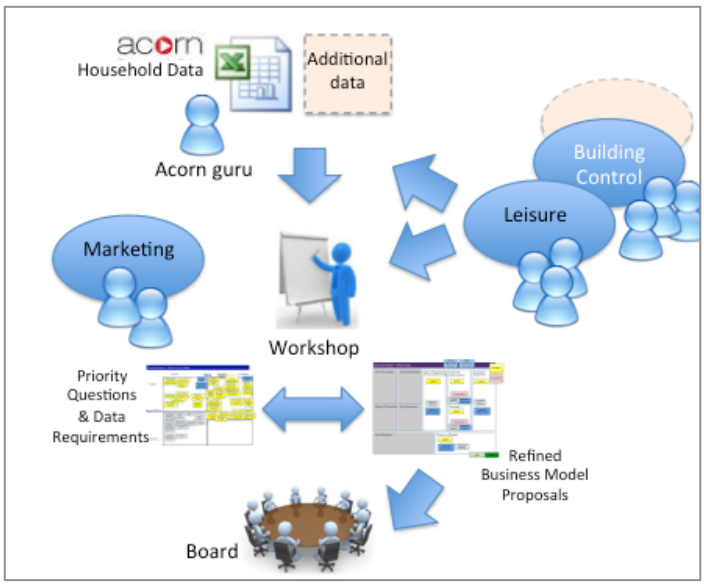


Figure 4-1 A visual summary of the GoCouncil Data effort

The market insight initiative is over and above the normal, ongoing requirements of coordinating marketing and Public Relations (PR) for the local council elections this year, internal communications newsletters and the like, let alone reacting to rather unpredictable news events and 'firefighting' councillor sniping or over-commitment of the council in the press. This leaves little bandwidth within the team to undertake significant tailored market research. Even so, many propositions will no doubt involve public consultation effort and inevitable involvement from her team before being able to launch them.

4.1.3 Research involvement in Market Insight Initiative

Only being in three days a week on a contracted basis, and with very little internal data capability or resources to help, Andrea contacted the Cranfield School of Management IS research team. GoCouncil's customer data-led approach was indeed of research interest. While Patricia took some convincing, and although Andrea had a concern she may want to micro-manage the initiative, the initial meeting went really well and I was on board.

My involvement kicked off with an onsite meeting with the wider project manager Albert, where it was quickly decided to develop a facilitated market insights workshop for functions, to prompt them to identify how they could use

Acorn data and market insights more generally to help them refine their service propositions. I also attended periodic project meetings held to monitor progress on market insights generally. These were typically organised around an action list maintained by Albert as overall transformation project manager.

I worked closely with Andrea to develop and test a workshop approach, on site once a week, working in a variety of meeting rooms, breakout areas and in the Marketing team's open plan office area. We ran a pilot with a small group of participants from the Leisure function, given that they had been chosen as a priority area for the business model initiative, as their cost challenge was greatest. It was also felt they had some familiarity with Acorn data compared to other groups.

The workshop included the following elements:

- An introduction to the Business Planning model and approach.
- An overview of Acorn data available (introducing Tanya as the Acorn Subject Matter Expert).
- An exercise to map proposed Leisure Revenue-generating Service propositions to wealthy target Acorn groups.
- The bulk of the workshop focused on facilitated and moderated exercises to generate market insight questions in relation to three selected propositions (one existing and two new ones) that were agreed with Leisure ahead of the workshop.

The exercise elements outlined aimed to get participants to use and practice the approaches, so that they could apply them more widely to the rest of their propositions to improve them, and also to identify where they may need assistance from Marketing to help them with obtaining further market insights. The exercise adopted the Business Canvas (Osterwalder & Pigneur: 2010) framework against which to generate and organise questions for each proposition, using post-it notes. After the workshop these questions were also organised into a portfolio of questions in relation to data available (or not). This data analysis is described in more detail in section 5.2 below. It was intended to

share these outputs with the Leisure leadership team, to facilitate prioritisation of subsequent market insight assistance.

It was also intended to roll the workshop out to all functions and for me to be involved with several more to assist and help other Marketing team members become confident to facilitate them unaided. To this end, Andrea and I refined and codified the workshop approach, based on a detailed debriefing of the first workshop. In the event, only another one was arranged with the Community Centre teams (working with Alex, another Marketing team member, as Andrea left the organisation in early May). Although initial planning discussions were held with the Planning Services team, a suitable date could not be settled on before my decision to withdraw from the field.

Helga, another experienced member of the Marketing team, had already been undertaking separate market insights work with the Planning Services team around some of their more advanced propositions, which they believed were ready for launching. This involved focus groups, rather than Acorn data led work, given that their services were mainly offered to intermediaries, e.g. builders and architects, rather than to households directly. Acorn data addressed households rather than businesses or such intermediaries. A similar 'flexing' in data focus was also evident for some of the Community Centre offers, e.g. divestment, as the target market for these was not likely to be households, although the selection of which Centres to potentially divest might be influenced by geographical spread of current users and related services.

While the results from the workshops were discussed with the Marketing team, I was very disappointed that the follow-up meetings, to discuss prioritising further market insight assistance, never occurred. This was initially because the main Leisure manager with whom we were working went off on a period of prolonged sick leave. Subsequently, this was overtaken by a shift in emphasis of the overall transformation project to address the increasing urgency to produce new business model service proposals, so that new or revamped services could be launched as soon as possible.

4.1.4 A shift to delivering Business Canvas Proposals

The initial pivot in this direction was evident very shortly after the first workshop, which received a very positive response and interest from senior managers in terms of increased profile for the wider project. Up until that time, much of the transformation effort appears to have been undertaken within a closed group of project team members with little wider visibility across the organisation. This prompted an invitation for Andrea and me to meet with the CEO to explain what we were doing. In the event it turned more into how we might be able help increase organisational visibility and engagement with the transformation effort, its related shift in staff mindset towards a more commercial and customer-centric view, as well as increased traction in terms of concrete service proposals launched.

This triggered an idea borrowed from previous programmes I'd seen during my prior consulting experience, of introducing a physical project office or 'war-room' to improve wider visibility of progress and proposals, and to facilitate more cross-functional engagement with function proposals under development. The idea immediately caught their imagination and was promptly implemented, although rebranded as an 'Engine Room' as more appropriate to the council culture, and perhaps also reflecting the new emphasis on producing business canvas artefacts.

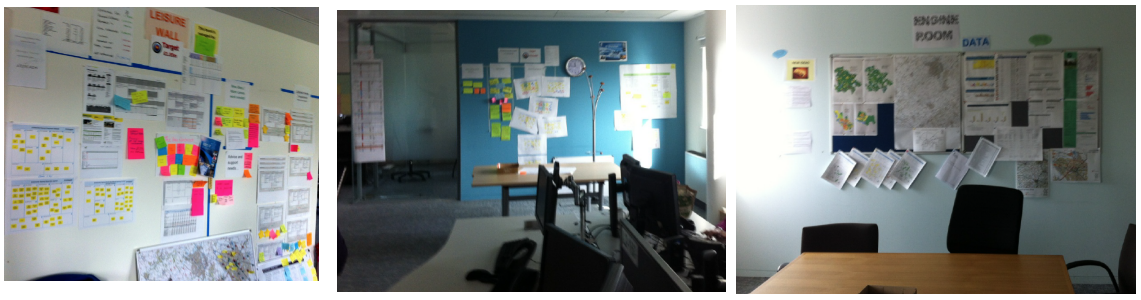


Figure 4-2 Engine Room with Function Wall Areas and Data Resource Wall

As depicted in these photos, the area design provided a wall to each function for their proposals under development, savings or revenue targets anticipated for each proposal or proposition (and progress to meet overall targets). A Resource wall was also created, with a collation of various elements of

potentially relevant data about the district, services and households (including the Acorn mappings for target affluent categories). As part of the resource wall, we tried to maintain a focus on grounding thinking in relevant market insights, together with basing a Marketing person in the Engine Room for teams to consult.

However, the effort to establish the Engine Room diverted Andrea and my attention and effort away from rolling out the market insights workshop facilitation and related follow-up in terms of market insight support. Senior management attention was now firmly engaged but centred on prioritising the production of business canvas proposals rather than exploratory market insights to inform or test the viability of propositions. This culminated in questions and increasing challenge from senior managers to clarify how much insight was enough to proceed with a proposal to the board. GoCouncil's waning interest in using and applying market insight to inform their business canvas proposals, prompted several discussions with my supervisor. This, together with other pragmatic considerations, led to a decision to withdraw from the field.

4.1.5 Other opportunistic meetings

While immersed at GoCouncil, an opportunity arose to visit a much larger council, referred to as OtherCouncil. GoCouncil identified them as more advanced in their use of market and customer data. I extensively interviewed their Customer Data Analytics champion, with a view to seeing what could be learned from them and reported back to Patricia.

I was also struck by the lack of IT department involvement and with Patricia's approval, had a meeting with the IT manager to better understand what customer data were currently captured in systems and to explore to what extent they would be able to support Patricia's initiative, at least in the medium term. This prompted a further, extended meeting, though this primarily focused on their perceived requirement for a CRM system.

4.2 Introducing InfraDig

4.2.1 Construction project background

InfraDig is a European public sector project organisation created specifically to build an extremely large, complex set of infrastructure, including nine sites and 42 kilometres of earthworks and tunnelling. This needs to be integrated with existing infrastructure and was planned to be built over several years. Its ultimate aim is to making service improvements for a large metropolitan environment, in terms of speed and integration.

As such it needs to address a large range of stakeholders' requirements. This includes the operating companies that will take over the running of the infrastructure and sites built (or upgraded), as well as the public authorities coordinating related services for the area and the general public. It operates as a design, commissioning and project delivery organisation, coordinating overall completion by various private sector sub-contracting construction and other companies, and handover to the ultimate operators.

It is a temporary organisation and will be wound up on completion of the project. The organisation comprises infrastructure construction engineers organised into design, project delivery and operational handover groups, with corporate support functions, including human resources and IT. There is widespread use of contractors and consultants within the project organisation, given its temporary nature.

4.2.2 Building an associated 'virtual' infrastructure

As part of the infrastructure project, InfraDig is committed contractually to build, collect and hand over an integrated set of data about the infrastructure constructed. As a public sector project, InfraDig also aspires to demonstrate good practice in terms of newly introduced Building Information Management (BIM) UK government requirements (BIM Industry Working Group, 2011). While BIM produces data aimed at facilitating smarter construction, it also aims to provide data in relation to the ultimate operators, to facilitate more effective

management and maintenance of the public infrastructure being built over its productive life. The expected productive life is estimated to be around 60 years.

This aspect of the project, i.e. producing data to improve infrastructure maintenance, characterised by InfraDig as creating a 'virtual' infrastructure, drew me to the project as an interesting research site.

> *"We're building two InfraDigs: we're building the physical InfraDig and the virtual InfraDig. And it's as simple as that. The loving care and attention that we pay to creating the physical world we should be giving the same love and attention to the virtual world because it's the virtual world that often gets used for managing and maintaining the physical world." (Chief Engineer)*

As an overall vision, it is compelling and sparks the imagination. It seems ambitious and futuristic, especially when illustrated using some three-dimensional prototype software – allowing a virtual tour of the infrastructure, enabling panels and layers to be removed and to access data related to elements pointed to or when hovering with a mouse pointer or trackpad.

Peering under the bonnet, to understand the mechanics of realising this vision, reveals a huge data collection effort, as illustrated in Figure 4-3.

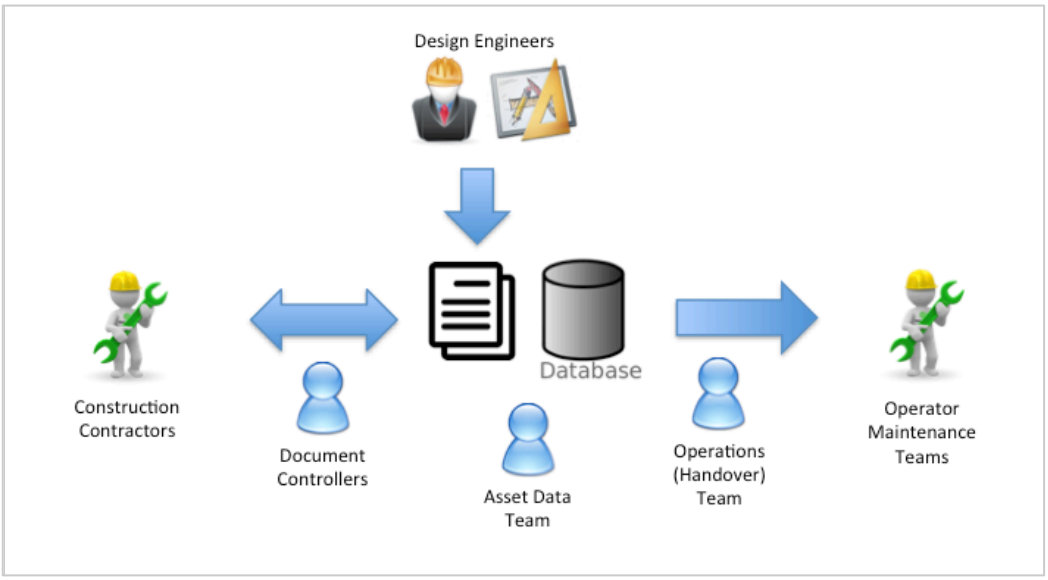


Figure 4-3 A visual summary of the InfraDig Data effort

In terms of scale, the technical design group of the project anticipates creating approximately 10m documents, 3m drawings, and 1.5m Computer Aided Design (CAD)[1] model files in relation to 2m assets or components.

Data collection builds on existing construction project practice, broadly involving InfraDig engineers designing and specifying the physical infrastructure to be built, captured in design documentation (CAD, drawings, component specifications, etc.). InfraDig used this design documentation as a basis for procuring contracts with various construction firms or contractors, who use these design documents to plan and execute construction.

InfraDig project managers and engineers assess delivery against these specifications before formal sign-offs and contract payment. 'As-built' changes are allowed for but have to be negotiated and agreed. Contractors are also responsible for delivering a final set of 'as-built' documentation, highlighting departures from the original design (called 'red-lining'), as well as operating and maintenance instructions for the assets or components delivered. The ultimate infrastructure operators are also involved, initially in signing off on specifications, and again when finally taking delivery, during a transition process captured in a 'handover plan'. The ultimate operators are also represented on the project itself, making up the majority of the Operations team responsible for formulating and executing the handover plan and transition process, and to promote knowledge transfer during the project.

Design documentation is maintained in a document repository, organised in a complex hierarchy related to functional units of infrastructure, which can be decomposed into constituent components, (e.g. an air conditioning system can be made up of various pump and other components). The InfraDig Asset Data team manages the document repository and its related structure, using a software package (eB) to keep track of documents and link to them (e.g. to the CAD system supplied by the same vendor). While hosted and supported by the

---

[1] **CAD** (**computer-aided design**) software is used by architects, engineers, drafters, artists, and others to create precision drawings or technical illustrations. It can be used to create two-dimensional (2-D) drawings or three-dimensional (3-D) models

IT department, using an outsourced provider, eB is primarily administered by a ‘super-user’ in the asset data team.

Data requirements are considered part of the specification and vary considerably for different types of equipment, so have been specified and signed off by class (or domain area) by the relevant infrastructure operator engineers, working with the InfraDig design engineers and asset data team. Some experienced infrastructure asset data management consultants have been retained by InfraDig to assist with this activity. On a day-to-day basis though, a distributed team of InfraDig document controllers collect documents and update the repository. They are co-located with the InfraDig project staff members, at the various construction sites, typically facing off against equivalent contractor staff supported by a small central team of document controllers, who also monitor this activity and document quality through periodic audits.

4.2.3 Research Involvement

I was introduced to the project by a data consultant to the project, via a British Computer Society Data Management interest group in which we were both involved. Based on a very positive introductory meeting, it was agreed that I participate in two projects within InfraDig over several months:

- **Asset KPI (Key Performance Indicator) project** – to specify requirements for a performance management system for the asset data collection effort described.
- **Information Management Strategy** – to develop a broader information management strategy for InfraDig.

An overview of the projects and my involvement in each is provided below.

Given the project nature of InfraDig and their considerable use of consultants and contractors, my part-time involvement fitted in very well with their culture and way of working, although I was not granted full email or intranet access, as were other contractors. While this sometimes represented a slight inconvenience, it didn’t impact on interaction, as I was regularly on site, could

email them from my university email and while working with Barbara, the asset data quality manager, more closely for some fieldwork, she was very helpful at tracking down and providing any documentation needed.

As a starting point, I was included on the asset data induction course, hosted at the eB supplier premises, where a 'sandbox' or test version was also made available for operators and others to explore and experiment. This was led jointly by InfraDig asset data and software supplier staff, and focused on the overall asset data programme, objectives, structure, and tooling that would be used. While onsite document controllers would undertake most data capture in the field, staff and contractors might need to interact with the software to obtain documents, etc. and they would be approached to complete and update relevant hard copy versions. The session provided me with a chance to chat informally to field staff and contractors attending about their views of the data initiative and how it compared to previous projects they had been involved with.

Asset KPI reporting project

This project aimed to define management information requirements, to facilitate tracking the progress and quality of the document collection activity described in 4.2.2 above. It represented a short, sharp project completed over a period of a month in February. The project aimed to produce a requirements specification, which would then be developed and implemented in due course, representing the first phase of a traditional project or 'waterfall' systems development lifecycle, which the consultancy adopted. A future implementation phase was planned but didn't occur during my period in the field.

A consultant project manager/business analyst was retained to manage the Asset KPI reporting project. He was from a large consultancy, which had been supporting the broader InfraDig since inception. He had supported them with another project to provide management reporting for a different aspect of the project, which had been well received by the chief technical engineer, who was sponsoring the project. The core of the project team comprised the asset data team members and their manager, as well as the asset data consultant they retained on a contract basis. The head of IT development attended several

meetings with an eye to integration with wider IT developments and systems, while the sponsor also sought to attend whenever possible, even if only for part of a meeting.

The main team activity was focused in five meetings, of which I attended four, combined with a review of documentation collated and presented by the consultant, e.g. draft requirements, datafields, mock-ups of reporting. He also undertook further investigation and interviews with one or two project managers, who oversee contract completion, who were intended to be major users of the management information produced to oversee their project contractors.

My main role during the project was as an observer at the meetings, although I did participate, contributing ideas and clarifications, as well as reviewing the draft design documents. However, my contribution was limited to general reporting advice, being unfamiliar with construction and related asset data as domain areas in my previous experience of IT governance consulting.

IM strategy

In addition to this central and fairly focused data initiative, I also participated in a broader initiative to develop an InfraDig Information Management (IM) strategy. The IM project lead for this initiative is an InfraDig manager, who also had overall responsibility for internal business project investment. He is also part of the IT executive team, so fairly senior, and the initiative is therefore also seen as an IT led project by other teams (e.g. Technical Design team).

As I joined, InfraDig seemed poised to retain a large consultancy to assist them, and initial discussions as to scope and focus for the project were underway, with the IT lead agreeing to pull together his thoughts in a brief. However, after one or two exploratory meetings things went quiet. Initially this was ostensibly due to the annual budget cycle intruding, and subsequently it was overtaken by an urgent need to address a significant operational failure with email systems (managed by a third party supplier), which also resulted in a change in CIO. Essentially no progress was made to formulate a wider IM strategy for InfraDig over the six month period I participated, nor per subsequent contact.

Against this background, I took the initiative, initially to chase the IM project lead for meetings. When the head of technical design engineering showed an interest in using CoP thinking to improve asset data strategy implementation, I suggested to Donald using the asset data initiative as a prioritised, bounded starting point for the wider IM strategy initiative. Asset Data certainly seemed likely to be a strategically important part of InfraDig's IM strategy. We discussed that this could then provide a template for formulating a wider IM strategy and approach.

We agreed that I would work with the asset data team to map key practice groups involved in the asset data initiative, related governance and other forums, policies, etc. with a view to identifying possible improvements. This mapping is described further in the Data Analysis, section 5.3, and culminated in a facilitated workshop, to explore ways to help ensure benefits sought from the Asset Data initiative might be realised. This work was undertaken, through interviews and meetings with various people involved in the asset data effort from the data team as well as other teams involved in the broader asset data effort. I worked closely with Barbara, a member of the Asset Data quality assurance team, who helped coordinate and attended all the meetings, knowing many of the individuals we sought to interview. Barbara disseminated the mapping internally with a view to encourage further refinement and extension of its use. She also helped coordinate a forum for Document Controllers, which I attended as part of my orienting effort and it proved a very useful insight into their perspective on the process.

As per discussion with my supervisor, we agreed that the debrief of the workshop, reporting back to the management team on the mapping, my related findings and recommendations for them to consider taking forward, provided a sensible and pragmatic point to withdraw from the field.

Additional opportunistic meetings

I met with the then CIO informally on two occasions, which provided a useful wider strategic context to IT within InfraDig and the wider construction industry. He alerted me to the BIM requirements and pointed me to related

documentation, which had not been mentioned by others interviewed until then. Indeed, I ended up disseminating this to Barbara, several asset data team members and wider interviewees.

The development manager and I also had a discussion about plans to address their requirement to handle and manage data generated by building movement sensors at city centre construction sites. Significant volumes of data are generated and a ‘big data’ tool, Qlickview, had been bought and was being proposed to address related reporting requirements. Unfortunately, no project and related research involvement flowed from this, although the discussion provided a useful exploration of data challenges posed by significant data volumes and the need for extremely quick responsiveness.

A meeting was also held with the member of the InfraDig operations team responsible for modelling movement through sites, based on the technical data collected and using a specialist related modelling tool. This provided an interesting insight into a further use of the asset data collected, as well as specialist modelling software, as distinct from the more generic eB asset management software used by the asset data team at InfraDig.

While at InfraDig, the opportunity also arose to see presentations by national infrastructure and metropolitan service asset data teams, regarding their data initiatives. The British Computer Society’s (BCS) Data Management Interest Group hosted both meetings.

5 DATA ANALYSIS APPROACH

5.1 Overview

Given the number of approaches adopted to analyse a range of different data, this chapter seeks to provide greater transparency about the process. The output and objective of ethnography is to produce a rich and compelling explanatory account of a social context or phenomenon, normally in the form of a monogram or thick description (Van Maanen: 2011, Hammersley & Atkinson: 2007). Given the intention to produce such an account for each case, my data analysis aimed to provide a firm and transparent grounding for these 'stories' (Van Maanen: 2011). With this in mind, the emphasis was on achieving a synthesis of the data collected rather than 'pattern-seeking' analysis and tabulation (Singh and Dickson: 2002, Miles and Huberman: 1994).

Figure 5-1 illustrates my iterative engagement with various kinds of data collected in order to make sense of what I observed and to distil these observations into theoretical contributions. The various forms of data analysis adopted and interaction with theory as part of my sensemaking are described in more detail below.

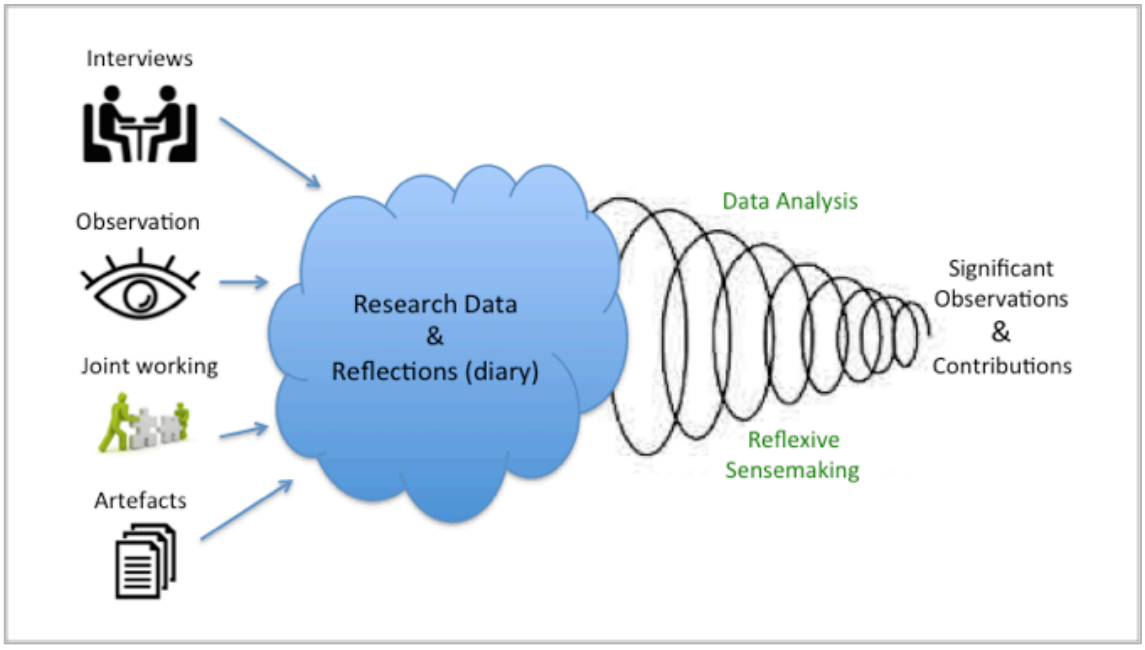


Figure 5-1 Visual overview of data analysis progression

Different strands of analysis

As illustrated above in Figure 5-1, various strands of analysis were undertaken from different starting points:

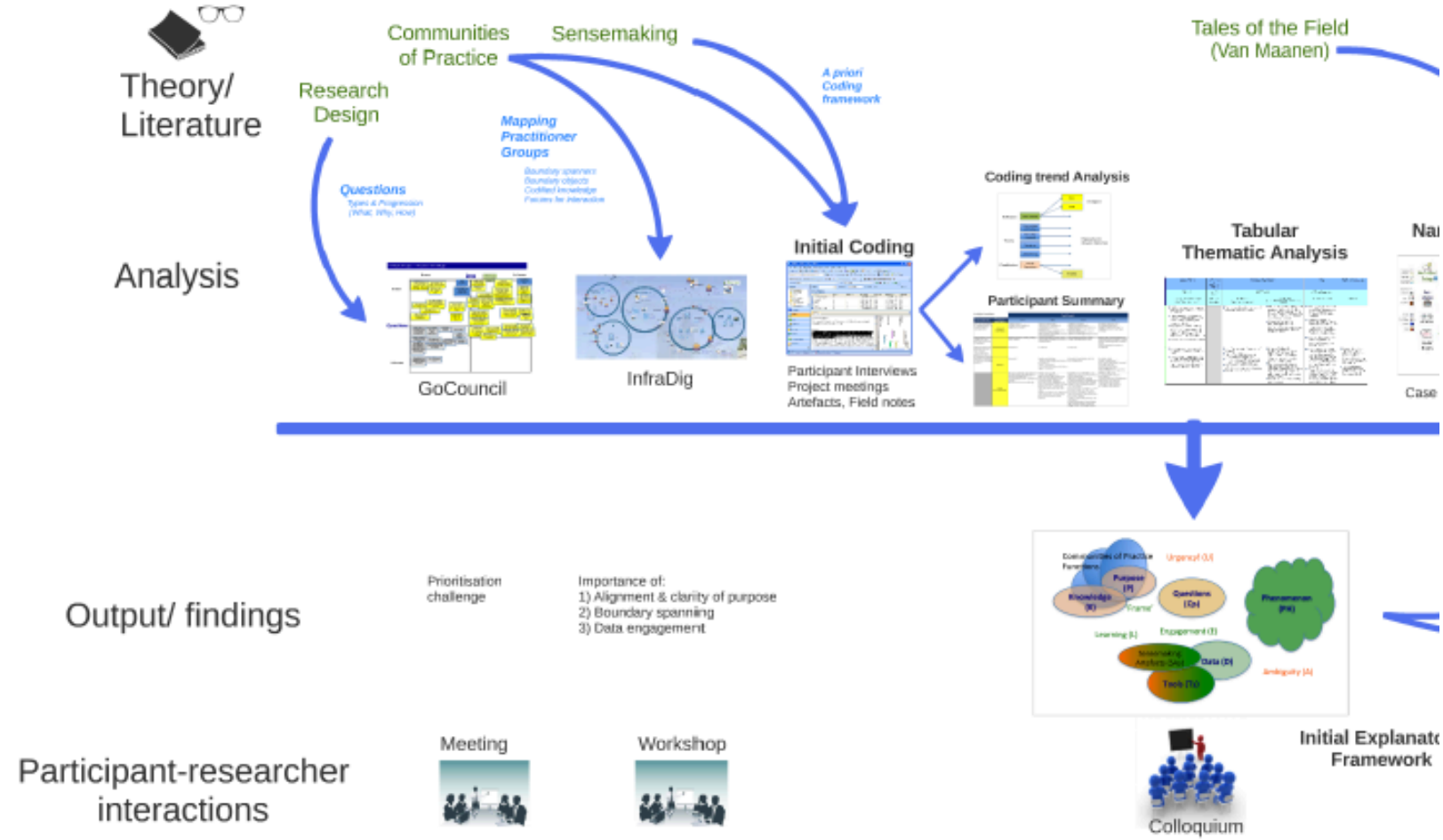
- Analysing workshop questions (GoCouncil)
- Mapping CoP (InfraDig)
- Coding (material collected, interviews, diary notes)
- Initial comparative thematic analysis
- Synthesis using a visual, narrative lens
- Reflexive cross-case comparative analysis

The first two strands of analysis reflect case-specific scenarios that emerged, which presented opportunities to collect particular kinds of data and undertake related forms of engagement with the data collected. The various strands also represent a general progression from specific, detailed analysis to broader forms of analysis aimed at synthesis.

Each of these is described in turn in the subsections below, followed by a note on reflexive and recursive aspects of data analysis. While some strands of analysis were specific to the case, others were cross-case in nature (e.g. analysing participant interview responses), and still others compared and contrasted cases or particular projects.

Dialectical role of theory

Consistent with an abductive research strategy, theory played both a sensitising and a creative dialectical role (Blaikie, 2010) during data analysis. This was most explicit in using Wenger's (1998) CoP framework as an *a priori* coding framework but was also influenced by continued reading, conferences and colloquia and conference paper feedback. This interaction is illustrated in Figure 5-2, in relation to various interim analysis stages and related artefacts produced during data analysis. Analytical artefacts indicated are touched on in the subsections below and samples and extracts are also included within Appendix D to provide further transparency.


Theory/
Literature
Communities
of Practice
Sensemaking
Research
Design
A priori
Coding
framework
Mapping
Practitioner
Groups
Questions
Tales of the Field
(Van Maanen)
Analysis
GoCouncil
InfraDig
Initial Coding
Participant Interviews
Project meetings
Artefacts, Field notes
Coding trend Analysis
Participant Summary
Tabular
Thematic Analysis
Output/ findings
Prioritisation
challenge
Importance of:
1) Alignment & clarity of purpose
2) Boundary spanning
3) Data engagement
Participant-researcher
interactions
Meeting
Workshop
Colloquium
Initial Explanat
Framework

(…\continued…)

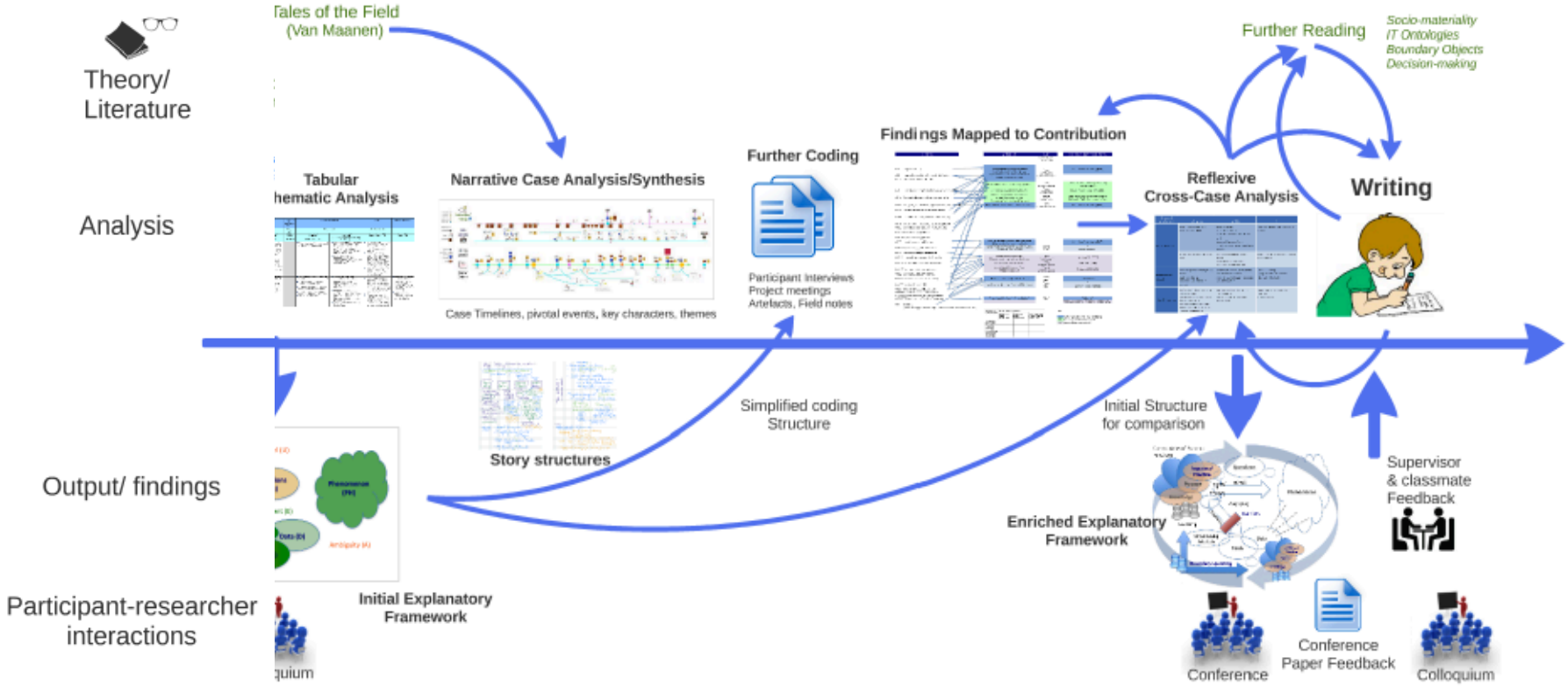


Figure 5-2 Overview of dialectical theory and other interaction my during data analysis

Alternate Timeline View

Data analysis occurred over an extended period, as a result of the considerable amount of data collected and having to summarise the bulk of recorded material myself. Analysis often alternated between cases and was punctuated by reporting on emerging findings at formal reviews, colloquia and through conference papers. This is depicted in an alternative calendar timeline view in Appendix A.3.

5.2 Question Analysis (GoCouncil)

The Marketing team organised several workshops to help functional areas consider and identify customer and market insight questions that were important to address. These questions related to new or enhanced service offerings they were considering, and needed to be addressed in order to be confident of the offerings' likely success and strengthen their case to the board for introducing them (and any associated investment required).

A Business Model Framework and Canvas[2] (as depicted in Figure 5-3) were already being introduced by the wider council transformation programme, with which functions could flesh out and present their new service offerings. Therefore, this was used as a basis for the workshop exercise, focusing on the customer and market side of the business model (i.e. the rest was shaded out). During the workshop, participants were asked to generate customer and market related questions on post-it notes and plot them on the related area of the business model. A facilitated discussion then reviewed, clarified, grouped and de-duplicated questions. The framework had been encountered by the CEO, based on his personal search and reading (Osterwalder & Pigneur: 2010).

[2] The reference used by GoCouncil refers to a business model as describing a rationale of how value is created, delivered and captured by an organisation The related canvas is proposed as a 'shared language' for describing, visualizing, assessing and changing such rationales (Osterwalder & Pigneur: 2010).

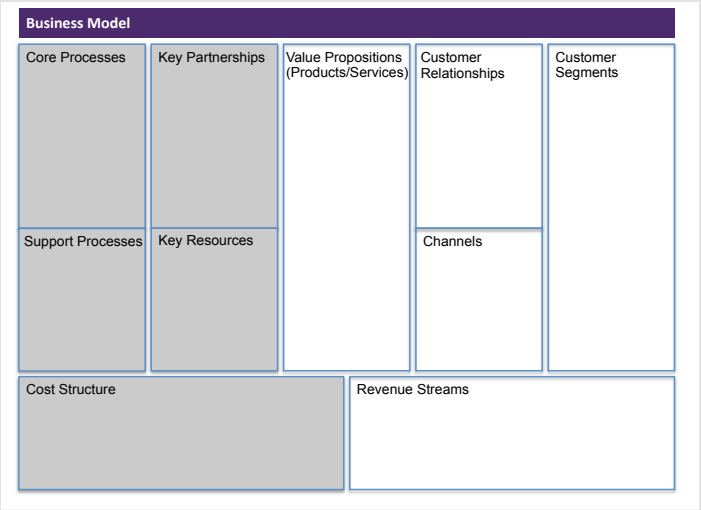


Figure 5-3 Business Model Canvas (based on Osterwalder & Pigneur: 2010: p.44)

Following the facilitated workshops, the questions generated were analysed for one service offer and shared with GoCouncil, with a view to a follow-up discussion with the Leisure team, firstly to demonstrate how they could complete similar analysis for remaining service offerings, and secondly use the analysis to help prioritise questions and related market insight research efforts, agree action plans, resource requirements and related logistics.

The analysis firstly included classifying and counting the potential questions generated, already mapped during the session onto the relevant section of the business model framework GoCouncil used to identify new revenue generating service offerings. Questions were classified into Blaikie's (2007) What, Why and How questions. Assumptions were 'surfaced' during the workshop and were similarly counted. By way of example, the post-it questions captured for the 'Theatre-in-the-village' service offer regarding Customer Relationships are shown in Figure 5-4 below. The two post-it questions in yellow were classified as What questions, while the two questions in grey were classified as requiring more clarification or refinement to isolate a clear market research question. The blue post-it represented a good improvement idea, in this case about a potential data source.

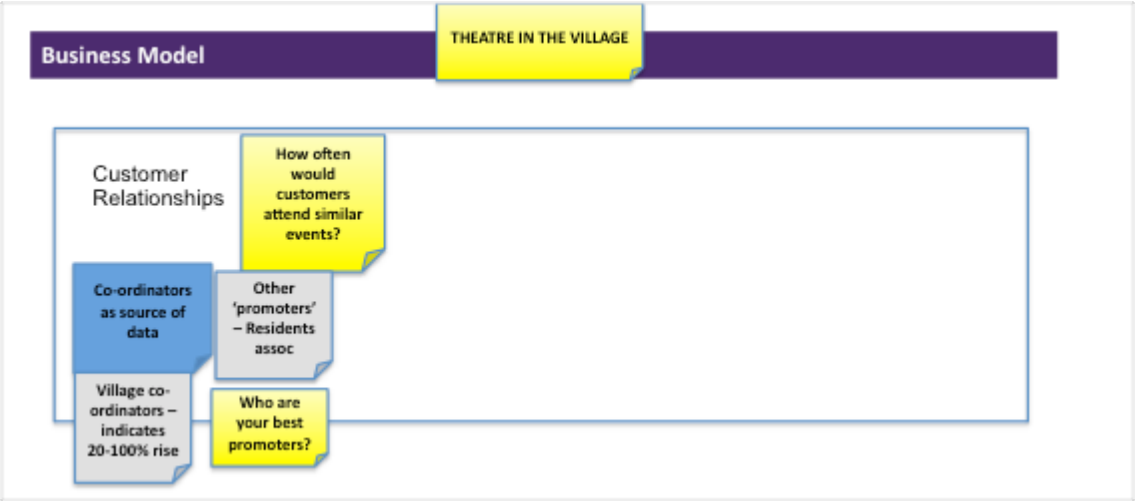


Figure 5-4 Sample extract of workshop Post-it Questions

Secondly, a portfolio of questions was generated for each service offering, plotting them on a two-by-two matrix, based on whether or not the data required to address the question were readily available (one dimension) and a preliminary assessment of the level of question clarity (the other dimension). This is shown in Figure 5-5. The two-by-two matrix and dimensions emerged out of initial supervisory discussions after my literature review, as a potentially useful portfolio analysis tool to aid grouping and prioritising questions and data analysis research efforts.

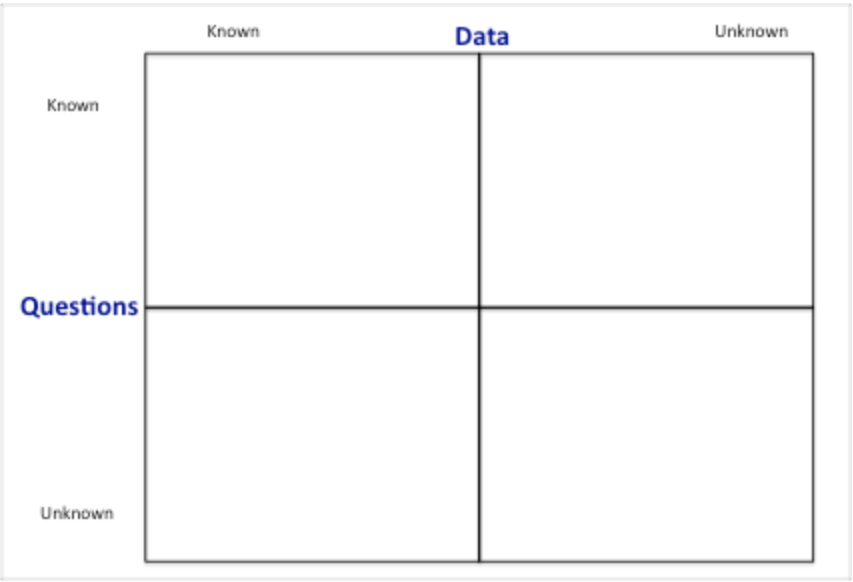


Figure 5-5 2x2 Question Portfolio-Prioritisation Framework

In terms of the sample extract introduced in Figure 5-4, the question: “Who are your best promoters?” was plotted in the top left quadrant of the 2x2 matrix as the data were available and the question was clear enough to Marketing to action. The two grey questions were provisionally mapped into the bottom left quadrant (but could on further clarification be migrated to another quadrant). The second question: “How often would customers attend similar events?” was plotted in the top right quadrant as data would need to be collected, e.g. through a customer survey in order to answer the question. The blue post-it in Figure 5-4 relates to a potential source of data rather than a question but was plotted for convenience and visibility in the top right quadrant, as these data still need to be collected. An illustrative set of both analyses for the Theatre-in-the-village offer is provided in Appendix D.1.

While the proposed follow-up meeting with Leisure didn’t occur, the analysis was completed for the remaining service offerings covered in the first workshop and also those for the second Community Centres workshop. A comparison of the results of the analysis for different service offerings and groups prompted reflection and several observations emerged which were shared with GoCouncil’s Marketing project team, when reviewing the workshop outputs with them. These were accepted and the only discussion points raised by GoCouncil related to very few re-classifications between Known and Unknown Data dimensions, given my limited knowledge of what data were and were not readily available. Related reflections and observations are included in an initial comparative thematic analysis outlined in section 5.5 (see also Appendix D.5).

5.3 Mapping people and groups (InfraDig)

While immersed at InfraDig, the opportunity arose to undertake an exercise to map the different groups and related members involved in the infrastructure asset data collection effort and stakeholder groups involved. This was done in conjunction with the wider information management strategy project. A visual mapping of the individuals, their roles, as well as formal and informal communication links and related forums, was created, using ideas from Wenger’s (1998) CoP framework, e.g. identifying different practitioner or

functional groupings, boundary spanners, codification of practice and boundary artefacts.

This exercise was done working closely with Barbara from InfraDig's data quality group and involved interviewing several other staff members to achieve a relatively complete view. Prezi (a flexible presentation tool) was used to capture, navigate and present the large amount of data collected, using icons, descriptive text and uploading related documents (e.g. terms of reference for forums, sample reporting). From InfraDig's perspective, the exercise was undertaken to get a better understanding of the key groups and individuals involved in collecting and using the infrastructure asset data, in order to locate and highlight potential communication and engagement gaps, issues and opportunities.

As an example of the mapping undertaken, an extract of the overall map is depicted in Figure 5-6. This captures onsite document data collection teams and roles on a typical sub-contractor project. While some details are not visible at this scale, the figure nevertheless provides an overview of the data collected and the mapping undertaken. The extract highlights the central structure at a sub-contractor project level and key data related contractor team members' mapping against InfraDig equivalents, coordinating requirements and monitoring data collection. Both teams use the same computer system, often collocated and collaborating, but with the formal relationship being governed contractually, with associated documents and formal meetings. Additional forums (e.g. periodic presentational update forums) are also indicated for coordination between contract focused InfraDig staff and Central or Head-Office staff. Several issues are flagged (e.g. related to eB system performance) using a visual fire icon in Figure 5-6. The involvement of further subcontractors is also shown, flagging the issue of a lack of visibility of sub-contractor data collection compliance and corrective action, etc. for InfraDig staff.

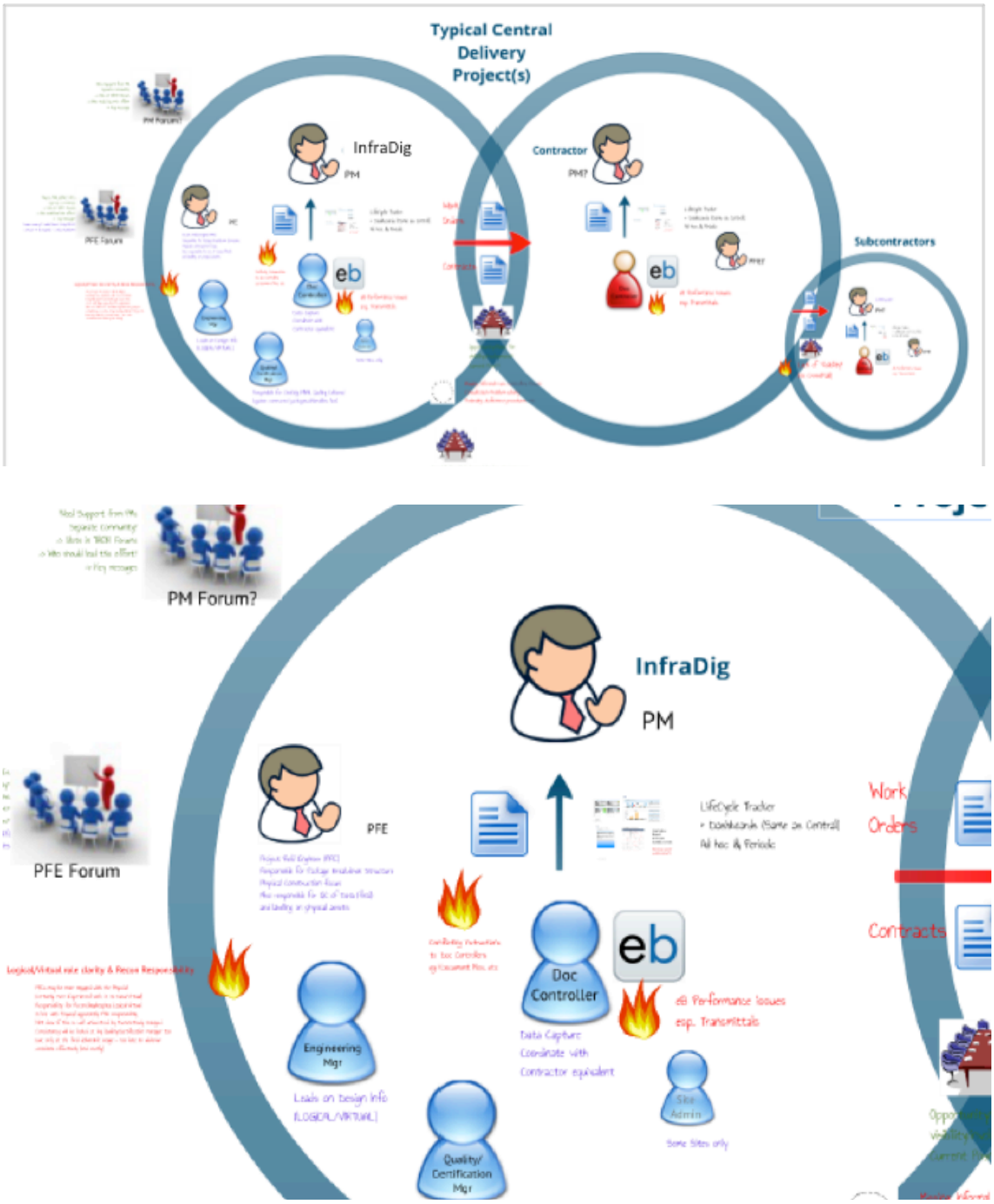


Figure 5-6 Sample Data Community Mapping extract and illustrative enlargement of one element

5.4 Coding using NVivo

Having left the field, I started by coding the meeting data collected within NVivo, the first and most granular level of analysis. This was done in conjunction with reviewing and summarising the interview and project meeting recordings, where available. My initial codes were based on Wenger's (1998) CoP situated learning framework, supplemented by Weick's (1995) sensemaking framework. Preliminary themes or initial field impressions, identified during research review discussions, were also coded for.

Coding occurred in two stages. The first was an intense period of fairly open and dense coding based on seven interviews, three project meetings, captured field notes and related artefacts collected, mainly drawn from the GoCouncil case. This is described further below. Even at this early stage of analysis, the main outlines and core components of the explanatory framework introduced in Chapter 7 started emerging. The second stage, for the remaining interviews, project meetings and artefacts, reflected a less dense coding, mainly based on the elements of the emerging framework, although still bearing in mind the CoP *a priori* codes and remaining open to additional themes and model refinements emerging.

An overview of the initial coding structure is shown on the left hand side in Figure 5-7. This illustrates how new codes emerged from preliminary coding before the structure stabilised. The review of initial coding also proved to be a useful source of reflection, highlighting areas of higher and lower coding density than expected, as well as triangulating coding density with initial themes identified and other strands of analysis. It was a source of reflexiveness too, noting what I was paying attention to and noticing, when engaged at a very detailed level of focus or analysis, e.g. on participant interactions during meetings.

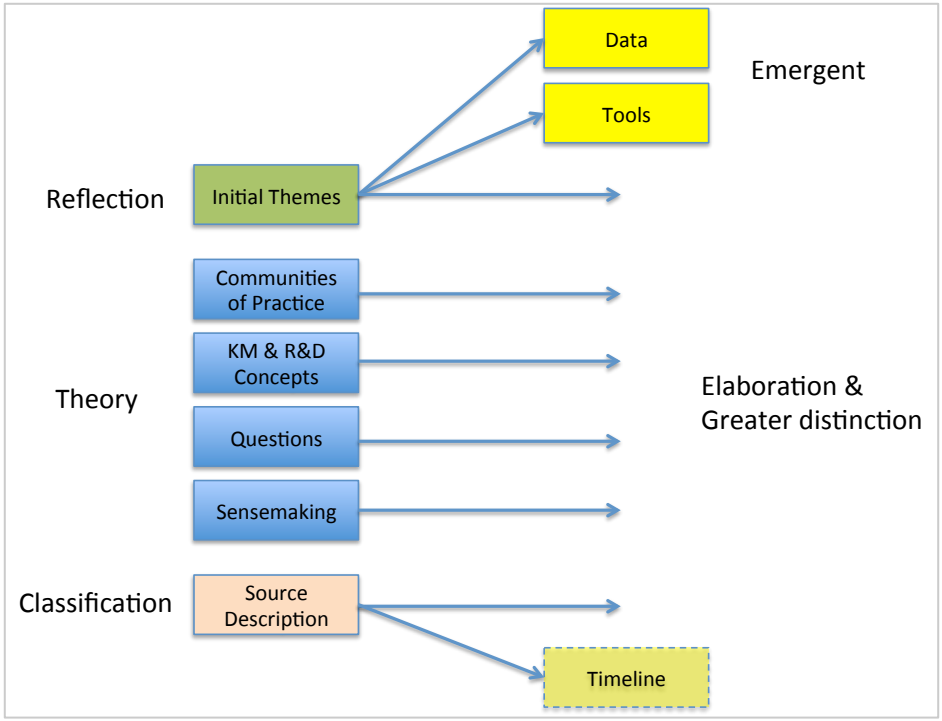


Figure 5-7 Initial (*a priori*) Coding Approach and Emerging Codes

5.4.1 Illustration of preliminary coding

By way of illustration, a short description is provided of how an in-depth participant interview was coded for the interview recording extract and review annotations shown in Figure 5-8 (an NVivo screenshot). The meeting recording and a scan of the associated meeting note were uploaded into NVivo. The recording was first coded against various source codes, such as case, participant name, meeting type (1:1 meeting), in order to facilitate tracking and later analysis. I then reviewed my related, scanned meeting note, before carefully listening to the audio recording, at the same time creating summary notes of the audio in NVivo (i.e. the audio was not fully transcribed in this instance). These notes were associated with particular sections of the audio recording to facilitate going back to listen to them again during subsequent iterations of coding and review.

As already indicated in the Methods chapter, recorded meetings and interviews were pragmatically fairly fully described and summarised rather than transcribed (Hammersley & Atkinson: 2007). Particularly interesting sections of the

recording were subsequently transcribed more fully, especially during writing up. For example, for pivotal scenes included in thick descriptions, these sections of audio were listened to again and transcribed more fully to ensure relevant details were reflected and related 'headnotes' refreshed (Schultze: 2000).

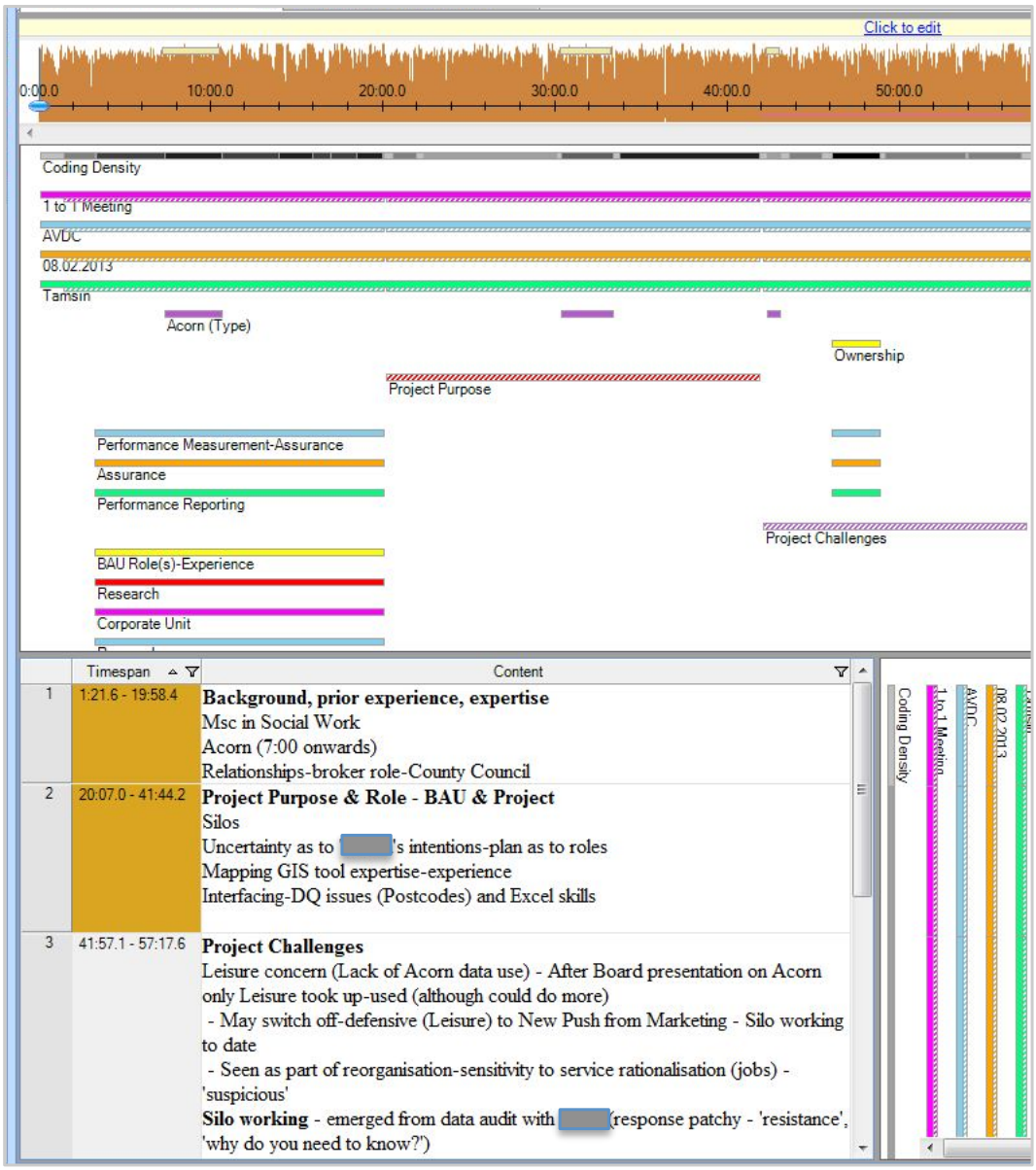


Figure 5-8 NVivo Extract of Interview Coding

Given the semi-structured nature of the in-depth participant interviews, the areas of the recording related to broad areas probed per the interview protocol developed, were coded as such, for example: Project Purpose, Project

Challenges, Approaches-Tools, participant Experience-Expertise and Project Role-Contribution (see also Appendix B). This can be seen in the coded extract in Figure 5-8. What is also evident from the extract is the creation and elaboration of sub-codes, for example, around types of experience and roles (e.g. Performance Measurement-Assurance, Research, Corporate unit).

This facilitated subsequent review and later comparison when producing the participant summary tabulation and comparison for participants. In addition to the coding informed by the interview protocol described above and interview responses, the audio recording was also coded using the *a priori* theoretical and thematic codes established (e.g. the ownership issue coded in the extract shown). Many of these coding approaches overlapped (e.g. the broker role identified in the extract summary was also an example of 'boundary spanning' in terms of the Community of Practice framework and was coded for both, although not visible in the extract shown).

This illustrates the sensitising role of these theoretical frameworks during coding and reflection, (although resulting in very dense coding, as can be seen from the coding density indication bar in the extract). Given the exploratory nature of preliminary coding, it was also very open in nature and new codes emerged and were elaborated (as can be seen from the Acorn coding in the extract shown, as a data type sub-code within the wider Data node, which emerged as a rich avenue for exploration and coding).

On completion of coding the audio recording and related notes, a short memo was captured in NVivo, to record my reflections on the interview, reflexiveness about interviewer and interviewee, and review and coding of the interview (e.g. new ideas and codes that emerged).

5.4.2 Analysis of preliminary coding

In view of the fairly dense and open coding approach adopted, after completing preliminary coding, the coding itself was reviewed. This was based on an

analysis, using Microsoft Excel™ (Excel)[3], of sequential code 'snapshots' (extracted from NVivo approximately every week or so). Per this analysis, the greatest changes to exploratory codes occurred early on; recognising that Data and Tools were important to code as nodes in their own right, these were split out from the overall initial themes identified. Beyond that, the broad nodal structure has remained with greater elaboration of related codes. This evolution is illustrated in Figure 5-7.

Coding density at this preliminary stage was relatively sparse across the coding structure, with the following nodes emerging as most densely coded (format: node: sub-node: sub-node), although this reflected the analysis of only a subset of data collected, mainly in relation to GoCouncil:

- Data: Issues-challenges-effort
- Initial Themes: Focal narrowing-Bounding: Functional-Experience
- Communities of Practice: CoP Learning Design: Engagement
- Function-Organisation: Silo Issue
- Questions: Explicit-Conscious
- Questions: Phenomenon Clarity: Customers and Services
- Sensemaking: Ambiguity-Uncertainty: How to proceed
- Sensemaking: Assumptions
- Sensemaking: Frames-Cues: Customer Orientation and Financial-Commercial Frame

Appendix D.3.1 provides a more detailed summary of coding instances at this preliminary stage, by group node and highlighting underlying elements or sub-nodes that were more densely coded.

As already highlighted earlier, the detailed coding evolution and density reviews outlined proved a useful source of reflection and reflexiveness, in terms of what I was paying attention to and noticing.

---

[3] While this particular supplier tool was used within cases as well, much of the later discussion about its use as a Tool is likely to apply equally to other spreadsheet solutions, i.e. denoting spreadsheets in general. The term Excel is used is used in the rest of the thesis rather than the full name.

5.4.3 Participant coding and comparative analysis

In terms of source coding, it was realised early on that the participant interview structure, as identified in the interview protocol, provides a useful nodal coding structure for sources.

This facilitated summarising participant interview responses and analysis in relation to particular areas of inquiry. An extract of the participant interview summaries produced is provided in Appendix D.4. This is presented as a columnar extension of the interview protocol, with the extracts focused on areas probed that are most relevant for the findings noted and related discussion.

5.4.4 Shift to summarising and direct coding for second phase

Given the growing frustration with NVivo, described further below, a more pragmatic approach was adopted for the remaining data. Remaining interviews and project meetings were reviewed by listening to the audio, transcribed or summarised, and the Word transcript coded directly using Comments. Reflections were noted at the end of each transcription or summarisation and subsequent coding.

This better suited the less dense coding during this phase and a growing emphasis on seeking synthesis rather than the pattern-seeking across interviews and meetings that NVivo seemed particularly suited to. I was not seeking patterns in my data but rather was seeking understanding.

5.4.5 Timeline view limitation

As coding progressed it became clear that it would be useful to organise my data along a time progression. It was envisaged this might facilitate an analysis of how participant (and my) sensemaking and reflection evolved longitudinally over the course of the case and analysis, and facilitate more narratively focused analysis, as described in section 5.6 below.

This was not straightforward in NVivo, and while a timeline source node was created and all data sources coded against it during my initial coding phase, it was still not clear how this could then be used to report a map or progressions

in data or coding over time, as originally hoped. An alternative approach was then identified, as described below, using a different tool.

5.5 Initial comparative thematic analysis

At the core of my initial data analysis, was a constant immersion in and engagement with my data (e.g. listening to and summarising meeting recordings), as well as iterations of reflection, from various starting points and paying attention to different cues and aspects of my data.

This process started as early as reflecting on my observations and capturing these thoughts in my field note journal; subsequently typing these up; and progressed to periodic reviews of these journal reflections ahead of meetings with my supervisor and a colloquium; through to the detailed coding and other analysis of the data collected; and during writing up.

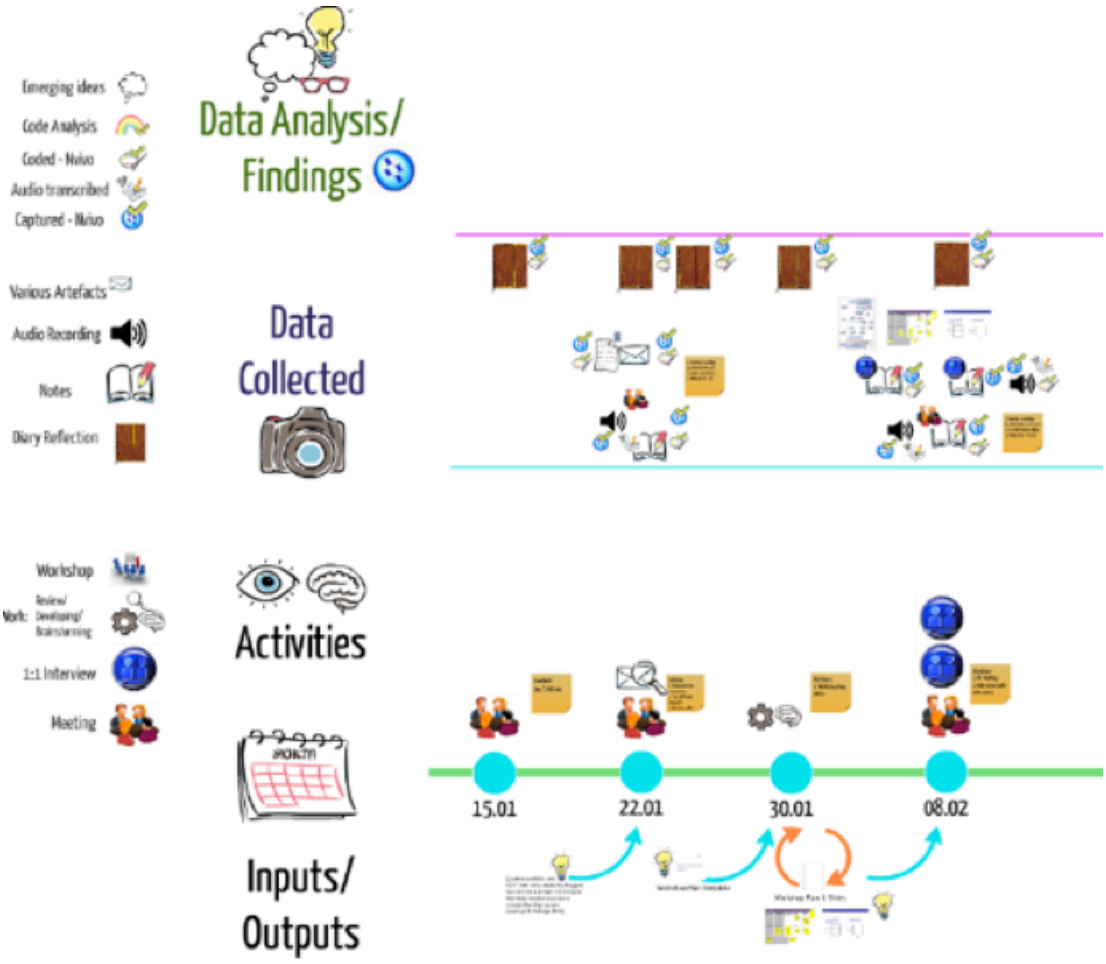
A comparison of the findings or sensemaking emerging from the different strands of my initial analysis was undertaken ahead of my second review. The comparative analysis is included in Appendix D.5. This is presented in tabular format with columns in rough timeline order from the left to the right. The findings were pragmatically grouped and themed using the high density coding group headings as a starting point, and then refined to include new themes or areas, as well as consolidating or realigning some findings where this seemed appropriate.

This exercise pointed to some consistency of findings, triangulating across different forms of analysis and reflection. It also highlighted how different forms of analysis brought different aspects of the phenomenon into focus, and how insights emerged over time.

5.6 Synthesis with visual and narrative lenses

The timeline constraint outlined, the need to progress towards a greater synthesis of my data, and the ultimate need to produce monograms or thick descriptions of my cases (Van Maanen: 2011), prompted me to also pursue narrative forms of analysis and engagement with my data, in parallel with the assessment of my preliminary coding effort.

As a starting point, Prezi was used as a canvas on which to plot the GoCouncil data collected along a clear timeline, based on my field notes. This provided me with a clear overview 'map' of my data, although with the ability to drill down to data elements collected, facilitating much easier visual engagement and reflection, while not losing sight of the data's wider context. A very short excerpt and key are shown in Figure 5 9 below to better explain the timeline approach adopted.



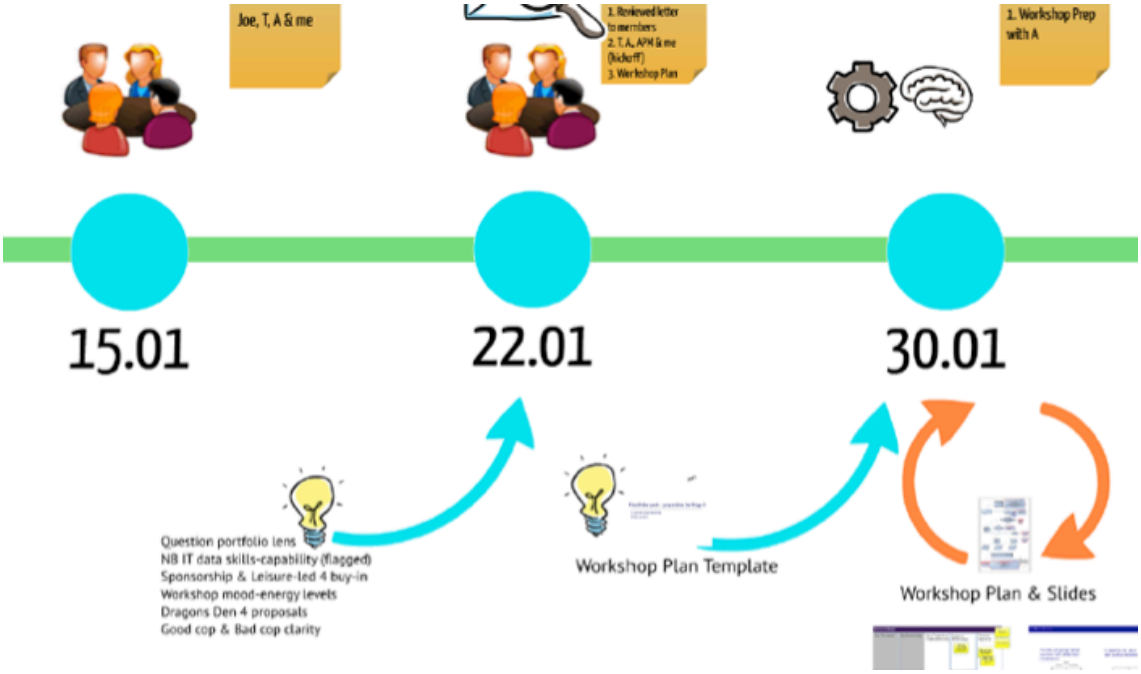


Figure 5-9 Narrative Timeline Map Extract with key and enlargement

The core and starting point for this mapping exercise was the case timeline (shown as a green line with blue dots and dates for days when I was on site, starting from the calendar icon). Above this timeline, for each day on site, I plotted icons for the various activities undertaken (together with a short post-it note with details (e.g. who was met with, etc.). As can be seen for the first three days, this involved a meeting on the 15th and 22nd of January, and reviewing some email project information on the 22nd, while I undertook some workshop preparation work on the 30th. Below the timeline, I recorded the contribution of any ideas or material (indicated by the lightbulbs) as well as some joint outputs (e.g. some of the workshop tools and templates indicated on the 30th January. Above the activities area (between the light blue and pink lines), I indicated the data collected during this period (e.g. diary entries each day, as well as a recording and meeting notes in relation to the meeting held on the 22nd January, and a copy of the email material). Details are captured in a small post-it note (for future reference) and the smaller surrounding icons track the capture of the data into NVivo and indicate that they have been coded. The area above the pink line is reserved for significant insights or ideas that emerge from analysis

or reflection over the course of the case, to highlight when they emerged and how key ideas or themes evolved also over the course of the case.

A similar mapping was undertaken for the InfraDig case. The complete visual timeline summaries produced for GoCouncil and InfraDig are included in Appendix D.6 (at an overview level rather than providing a view of details captured).

These visual mappings facilitated focusing on discerning and constructing a narrative synthesis for each case, allowing me to identify key episodes and turning points. Based on this reflection and engagement, a narrative outline structure was brainstormed for each case, as a starting point for monograms or thick descriptions for each, identifying key themes and features I wanted to draw attention to, as well as related episodes illustrating these. These initial brainstorms are included for reference in Appendix D.7.

I also pursued descriptive writing up of the cases' findings, in the first instance for GoCouncil for my second panel review, and in the case of InfraDig for a conference paper presented (Douglas & Peppard: 2014), as well as a further conference paper, generating useful feedback. For my third review, an initial journal article was drafted, based on the InfraDig paper referenced above. However, during a review with my co-supervisors, it was decided to draft a further article based on both cases for my third review. These writing efforts have evolved with feedback received and further reflection, representing an interesting dialectic device to complement my other analysis, at the same time also offering an opportunity to practice writing. Particular conference themes also offered a useful opportunity to practice looking at my data and arguments through the lens of a particular research question, with which to focus my writing and argument for a particular audience and purpose. The ultimate monograms or thick descriptions that emerged are included in Chapter 6.

This creative design engagement with my data brought into focus several filtering and prioritisation choices, forcing me to limit inclusions to the most important explanatory facets of the phenomenon and incidents in relation to the research question posed. It also revealed several implicit assumptions, which

could then be made available for reflection and usefully triangulated against emerging findings from coding and other forms of data analysis.

5.7 A reflexive cross-case comparison

Having completed the data analysis for each case independently, a reflexive cross-case analysis was also undertaken, treating my research as a further, recursive case of trying to generate insight from data. This seemed to offer a unique opportunity for triangulation and reflexiveness, given the nature of my own research, trying to make sense of and generate insights from my data.

The initial explanatory framework emerging from my earlier data analysis and reflection was used as a basic structure for this comparison, informed by my ongoing literature engagement and writing efforts. Based on these it was also aligned to facilitate a link to the three contribution areas emerging, though keeping an open mind for new ideas. Indeed the idea of barriers to engagement, different levels of interaction afforded across different kinds of boundaries and the possibility of different types of software affording different kinds of mediation and facilitation of data engagement, all emerged from this analysis.

The cross-case comparison is presented in Chapter 7, following a brief introduction of the explanatory framework for data projects emerging from my analysis. The cross-case comparison is thus also used as an illustration of the explanatory model.

5.8 Iterative refinement

Ethnographic data analysis "is a *reflective* and *iterative* process of identifying categories developed from extracts of the various data sources" (italics added), according to Singh and Dickson (2002: p.127). The various strands and methods adopted and described have indeed facilitated and aided reflection, through iterative engagement with the data collected, from various starting points, levels of analysis and foci for analysis, illuminating different aspects of the phenomenon and offering opportunities for triangulation and reflexiveness (Davies: 2008, Singh & Dickson: 2002, Eden & Huxham: 2002). Together with

my ongoing engagement with theory through reading the literature and feedback, as illustrated in Figure 5-2, this helped refine the emerging explanatory framework presented in Chapter 7, as a starting point for the cross-case comparison.

One example of such refinement, in addition to those already mentioned, is the recognition of the role of participant theories-in-action, which remained mostly implicit. Successive reflection and discussion isolated two different logics or theories at play in the projects, relating to their inquiry and the role of data, on the one hand, versus how any answers and related data were likely to be used to realise benefits, on the other, even though these logics were sometimes combined.

This example highlights the progressive clarity and specificity that emerges and is a hallmark of the ethnographic approach across its various stages, from inception, through literature review, fieldwork, data analysis and writing up (Hammersley & Atkinson: 2007). This can perhaps best be thought of as a spiral of increasing clarity, as illustrated in terms of my research in Figure 5-10. This progression continued throughout crafting and drafting of the ethnographic monograms outlined, reporting on emerging findings and in the writing of this thesis, though not as neatly or sequentially as implied in Figure 5-10.

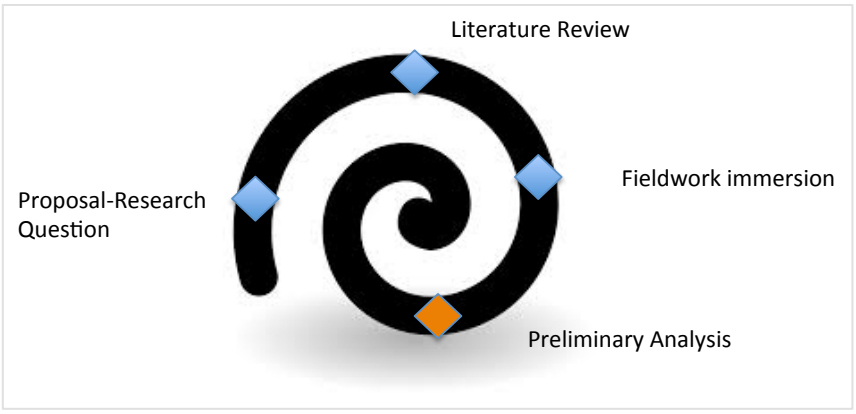


Figure 5-10 Ethnographic spiral of progressive clarity

This progression towards greater clarity and specificity, often via problematising and apparent ambiguity is, to some degree, also reflected in the structure of the

remainder of the thesis. First, descriptions of the two cases are provided including rich, episodic and confessional elements in Chapter 6. Second, an overall emerging explanatory framework is introduced in Chapter 7, based on the data analysed and iterative reflection. This is illustrated using the two cases, and used as a structure for a cross-case comparison, using my own research as a reflexive third case. The chapter concludes with a comparative cross-case summary and an enriched explanatory model. In Chapter 8, I then discuss some of the key themes emerging in relation to the research question and wider literature, before considering further avenues for inquiry, some practice implications and some limitations.

6 THICK DESCRIPTIONS OF CASES

In this section, two accounts are presented of the contrasting cases studied. For different reasons, neither generated much insight, nor significant related benefits. Indeed, the case descriptions highlight various challenges faced by participants pursuing insight. The accounts represent rich, episodic and confessional descriptions of what was observed (Van Maanen: 2011), which explicitly acknowledge my role and perspective as researcher. Both of these descriptions are based on the data analysis outlined in Chapter 5, especially the visual and literary analysis described in section 5.6.

GoCouncil – This account centres on a particularly challenging and pivotal project meeting some way into the ethnographic research involvement. This allows some perspective on progress, an insight into a data project meeting, as well as explaining the data led rationale behind the market insight initiative, as a new sponsor comes on board and joins the project.

InfraDig – Given a far more complex organisational setting, with research involvement providing a partial, early-stage view of a much larger, multi-year data initiative, this account focuses on a reflective look-back at two data projects shadowed, in the lead-up to a benefits workshop debrief meeting at the end of my involvement.

6.1 Market insights for a new world (GoCouncil)

Two months in and it seems Patricia has positively warmed to our ideas of how to structure the insight workshops for various teams that she had in mind. These sessions are intended to introduce the Acorn data, encourage engagement with it and position Marketing's ongoing involvement to ensure propositions are grounded in appropriate market insight. They will be rolled out to each functional team in succession. The first pilot session has gone down very well with the Leisure team, in spite of their feeling pressurised from the CEO to come up with concrete proposals and a concern they may be a bit defensive. My long experience of running workshops has helped, as well as my independence, prompting commercial questions, which certainly got them

thinking... Lots of good market insight questions and related data requirements have been identified and a new challenge is going to be prioritising them. While my 2x2 analysis of questions and available data seems useful, we still need to debrief with Leisure to prioritise and agree follow-on market insight actions and who will do what. Leisure are proving elusive and the workshop follow-up meeting hasn't been pinned down yet…

In the meantime, news has got back to the CEO and Patricia is getting positive recognition for Marketing's proactive coaching at recent board meetings. Still, progress is too slow and the CEO is getting increasingly impatient at the lack of business model proposals being presented to the board.

****

With this background, we turn to an account of a pivotal meeting, aimed at fleshing out an idea to energise and generate more business model proposals...

6.1.1 An 'Engine Room' for more business models...

This meeting is crucial to get the Engine Room idea off the ground and increase project momentum. Unfortunately Patricia, the head of Marketing cannot be here to lead the meeting. She is receiving a long service award this afternoon, at a special ceremony, but she is happy for Andrea and me to go ahead with the meeting to avoid any delay in launching the Engine Room. Having secured the deputy CEO, Justin, to sponsor the initiative, lining up diaries with him to arrange the meeting has been a *real* challenge.

The rest of the market insight team are all present, as well as Albert, the wider transformation project manager, sensing the meeting's importance. Really this should have been his idea. After all we are supposed to be *supporting* the transformation initiative with market insight to inform new propositions and market test them.

Albert is calling the meeting to order...

Andrea and Albert have already presented the broad idea to the board a week ago, so today's meeting is aimed at detailed planning and logistics to get the

Engine Room up and running. Patricia had rightly insisted Andrea attend the board meeting to present the idea, and also help rework the slide deck Albert had come up with, to make sure it had the right impact on the board. She knew Andrea would present the idea well, with flair and conviction, based on our earlier brainstormed ideas. Patricia was concerned Albert's pack was too wooden and boring, full of management or project-speak, so wouldn't really grab the board's imagination or have the right impact, presenting it more as a project management solution. Andrea and I agreed. Andrea's quiet, almost gentle manner belies a sussed, political operator with a long campaigning background. She combines a refreshingly direct approach, almost fearless, with an understanding of how to build support and keep key players like Albert on side and on track.

The pack had lost some of the visual richness we had included in our pitch to the bi-monthly market insight project meeting a week or two before. We had brainstormed the idea of a 'war room', based on a very successful one I'd come across at a big Telecoms firm a few years ago for their key programmes office. We mocked up some rich picture sketches of ideas for the Engine Room walls, to provide some tangible visualisation of what it might look like. We took great care to craft colourful, impactful pipeline charts that at a glance show progress, can be easily understood by anyone popping in to the 'war room', and also encourage participating, contributing ideas and questions using post-it notes. We were determined it had to be fun, engaging and yet focus attention on critical success factors, progress and generating viable propositions, grounded in appropriate market insight.

As some of the excerpts in Figure 6-1 show, we had left them purposely unfinished in style rather than as 'fait accompli answers'. The intention was to provide 'straw-men' to spark ideas, creativity and evolve something they felt joint ownership for through wider personal emotional and creative investment. We envisaged wall areas for each functional team addressed, e.g. Leisure, as well as Marketing and Finance (to provide data resources, make/prioritise requests and book coaching sessions), and the overall transformation project

team (e.g. overall status, propositions, savings/revenue achieved, etc.). These are reflected in the three examples below.

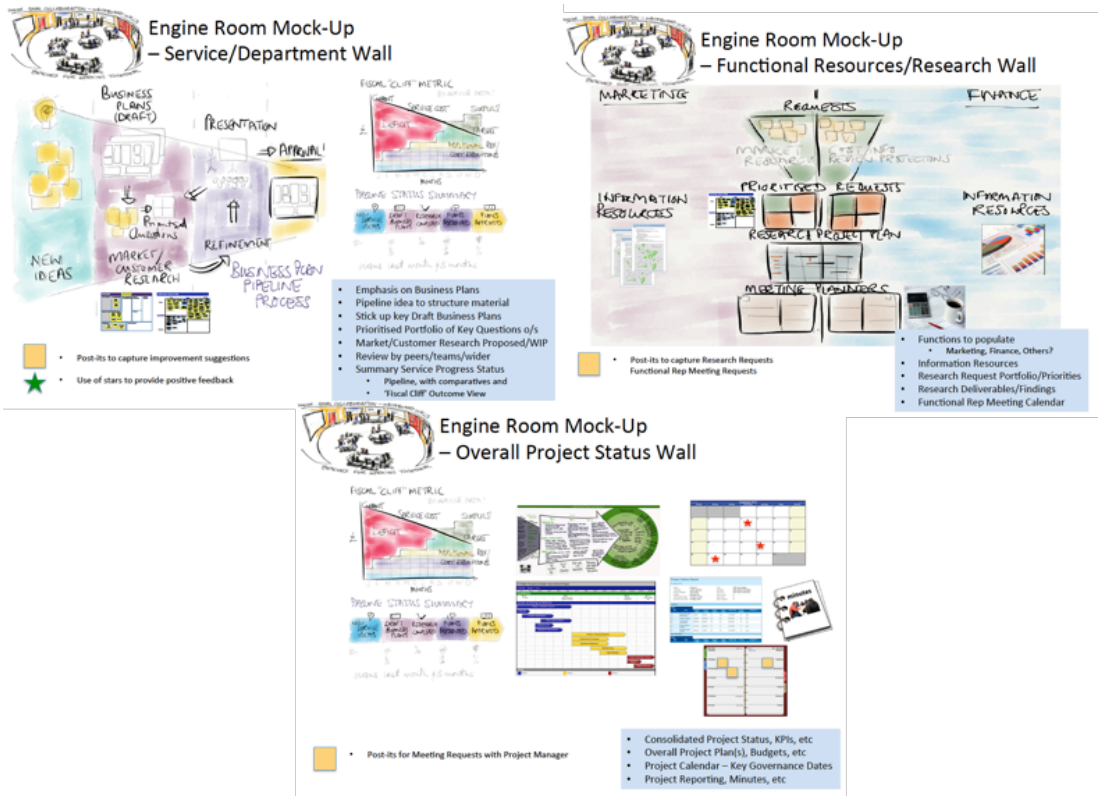


Figure 6-1 Engine Room Wall Mock-ups

Patricia and the rest of the market insights team had immediately grasped and liked the idea, generating lots of enthusiastic discussion. Initially this focused on renaming the ‘war room’ something more positive and less martial and threatening (given the amount of staff sensitivity about the transformation). They fairly quickly settled on Engine Room as a more palatable metaphor and name for it. The discussion quickly progressed to the practicalities of where best to locate it, to achieve the best visibility and widest engagement, which rapidly became mired in the politics of freeing up and securing the best space.

This brings us back to the need for securing an appropriate sponsor to cut through some of these hurdles and the detailed planning focus for today’s market insight project meeting.

Without any preamble, introduction of Justin, or framing of the purpose and focus for the meeting though, Albert jumps straight to his key action items list,

carried over from the last market insight project meeting. We typically use this as a default agenda for our meetings. He turns to Tanya for an update on where we are on producing and circulating Acorn factsheet summaries for the target household groups.

Tanya explains that she has some final checking to do but that they will be finalised later today. Tanya is our resident Acorn guru and has been presenting the Acorn introduction and overview slot at the market insight workshops. Based in the corporate team, she compiles periodic council performance submissions to central government and performs related assessments and audits. She also previously pulled together the district summary information pack, a really useful repository of information, which she thinks is underused. Tanya feels strongly that GoCouncil could be more data driven, coming from a social services background, which included research training. She was taken aback when she first arrived at GoCouncil, at their relative lack of data use and evidence-based approach to decision-making. This initiative seemed a good opportunity to improve the use of the considerable data the council had. Patricia did well to secure her involvement, albeit on a part-time basis, as she is the person most familiar with Acorn within the council and generally a fund of information – indeed the 'go-to' person for most teams after some data or information. She also manages the Acorn data subscription, liaising and working closely with the county council Acorn data team too, to coordinate and disseminate updated data and provide more useful Excel categorisations that are consistent across councils within the wider county.

Justin challenges our Acorn group prioritisation

For Justin's benefit, Albert starts providing some background on how we are summarising the Acorn data into factsheets for department teams to use to inform their new business propositions but Justin interjects wanting to know "*in plain English, what does this mean in terms of bringing forward a product to green light?*" His tone pulls everyone up short....

Tanya begins to explain that these represent the top five Acorn household categories but Justin interjects again, asking about the basis of the categorisation and what products we are planning to launch to them...

Tanya hesitates, floundering...

Meanwhile I'm wondering, surely this approach had been discussed and agreed earlier by Patricia with the board? I'm suddenly trying to remember if this was included in the related board proposal briefing on the market insight initiative I saw a while back...

Meanwhile Albert steps in and asks if Andrea has the market insights background approach material to hand. She wasn't expecting to have to brief Justin on this at the meeting though, so she hasn't got it with her but goes on to explain Patricia's rationale. It's not about products it's about knowing who the customers are, where they live, then mapping them in the vale and building up profiles, so that when we go into the Engine Room workshops, we want to be able to say to people: these are the top five and of your offers, which of your offers match these groups. Or which new offers can we match these groups to? So instead of producing another dance for disabled people, they produce a premier service for wealthy mature professionals who can pay for it, if that makes sense...

However, Justin challenges again that in terms of value propositions, aren't we limiting them to those propositions, which are just interesting to these subsections?

This sparks a lively, if slightly defensive, discussion. Andrea argues that we don't currently sell anything to these groups, an opportunity missed given their residual income. Customer profile data are not routinely collected for services currently and Tanya adds that this is where the Acorn data can provide useful information about people's interests. Andrea agrees, highlighting the Theatre-in-the-village Leisure proposition as an example: we mapped the wealthiest groups and discovered there are wealthy areas where we don't put on any theatre while we were for poorer areas; then challenging Leisure to get more

money by targeting these wealthier groups, trying to steer them towards looking at the data more when they are planning their services.

She and I also touch on the absence of systems and a data team, compared to other councils, to look at data across various service silos, hindering us obtaining more fully formed images of our customers and getting to know our customers better, highlighting that this aspect has been 'parked' for now but does need addressing too.

After some discussion, Justin restates his challenge though – why are we only targeting these groups? He cites the example of the Galaxy S4 phone sales distribution occurring in unexpected areas because different groups prioritise different spending, also pointing out some lower income groups have much better trainers than he'll ever have.

The meeting increasingly feels like an interrogation or audit of the market insights approach being taken, rather than focusing on how to get the Engine Room idea off the starting blocks. I am taken aback by Justin's continued aggressive and challenging tone. He seems to be going straight for the jugular, intensely focused on the urgency of getting to product launch outcomes and when Patricia is not present to explain or defend her thinking and approach...

In the interests of trying to refocus the meeting towards a discussion of the Engine Room, my consulting training kicks in, and I carefully acknowledge that Justin's challenge is a good one and that we can look at producing profiles for other groups too. Tanya agrees and suggests perhaps doing so based on the new version of Acorn data due out soon. Before I can steer things back towards the Engine Room discussion though, the careering steam train of a discussion is off again...

Justin weighs in with a second concern, that we should gear our use of Acorn data around the products we think are nearing production, and asking which these are. Albert gives him an update on which teams have been through the workshop process so far and which are scheduled to next, stressing that we

have been helping people start getting into the idea that customer analysis needs to start to inform their thinking, which Justin acknowledges is important.

In response to Justin's concern to gear Acorn data use, Tanya points out that there are several things you can do with the Acorn data, and that the data can also prompt ideas. *As I've been doing the group profiles, I've been looking at it and going... ooh, gardening, gardening... 50% of every group list gardening as a hobby. So I've come up with this idea, which I've been discussing with Leisure.* Albert agrees, giving another example of a line about culture coming out in a workshop in a similar way to gardening, pointing out that this is the relationship between value propositions and customer, and you can start with either. Andrea and I concur.

I also point out that the Acorn data and business canvas model, both used in our workshops, don't prompt participants to consider 'market analysis'. This could also be addressed though, by combining and extending Acorn data with other data – for instance, mapping where all the existing garden retailers are combined with the related Acorn data to identify areas of potential over- and under-supply.

This prompts Justin to clarify that what he is challenging is, do we know how to use these data, or do we just chuck the data in the air and see what falls? Andrea acknowledges that she only knows how to a degree, that when she got here, I knew nothing about insight and have learned through the project. She and I point out that having spoken to OtherCouncil, they are much more data led than we are as a council, with a dedicated team to work with their service areas and more integrated data across service areas to allow for wider analysis, though their data analysis is focused on how best to enable self service delivery of their services rather than generating revenue.

A concern about analysis paralysis

Justin reiterates his requirement for a clear link between the data analysis we do and how we are going to use it. Just doing data analysis because it is a nice to have or interesting... before you start research you need to know what research question you are trying to answer.

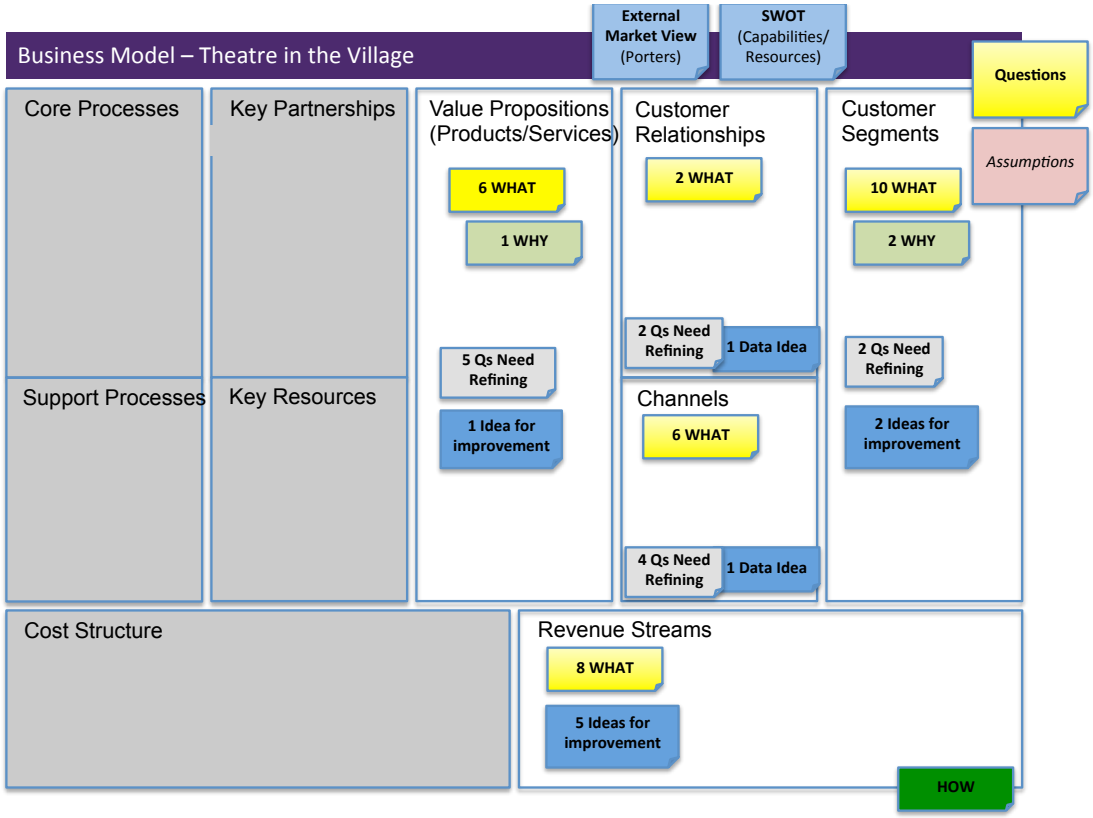
This reveals a concern that we might be doing too much or unnecessary data analysis, so I emphasise that no significant data analysis has in fact been undertaken, re-emphasising Albert's earlier point. Rather, we have been offering teams the Acorn data as an aid to their thinking. So far, we have only mapped existing services to Acorn categories for workshop purposes. Albert also points to Tanya developing a bit of a self-help guide for Acorn data. Trying to reassure Justin, I explain that our efforts so far have been focused on helping teams identify their key market and customer questions and data requirements in relation to their propositions. Andrea and I go on to explain how we've done this in workshops using the business canvas for one or two propositions so that the teams could then do this for themselves for other propositions they were revisiting or developing.

However, this doesn't reassure Justin, who now testily restates his concern to focus any data analysis. He fears that we are doing analysis because it is interesting. We can always ask more questions before the green light... the point is this organisation is paralysed by indecision because it likes to analyse problems to death, so what I want is not necessarily more data and more questions, I want to be very clear when we draw the line, saying that's as much information as we need to make a decision.

His tone and tenor are getting my back up at this point, as our concern within the project team has been the relative *lack* of any customer data led thinking, not too much analysis. Masking my rising irritation, I emphasise that there is no intention to do any analysis that hasn't gone through Patricia and repeat that very little data analysis has been done to date – as far as I'm aware only a focus group exercise to test building control's new consulting proposition. Acknowledging Justin's concern as valid, Albert adds that the only other

analysis has been an exercise on pricing for providing outsourced payroll administration services. I confirm that Patricia, based on our conversations with her, is likely to ask pointed questions when business proposals are presented to the board in due course, about customer assumptions underpinning them and the extent to which these have been tested or represent a gut feel.

Andrea intervenes and wonders out loud if showing Justin some question output from our workshop exercise for a proposition and question analysis would be helpful. Luckily, I have these to hand and produce them for Justin to take a look at (see Figure 6-2 for an overview snapshot for one proposition and Appendix D.1 for the detail).



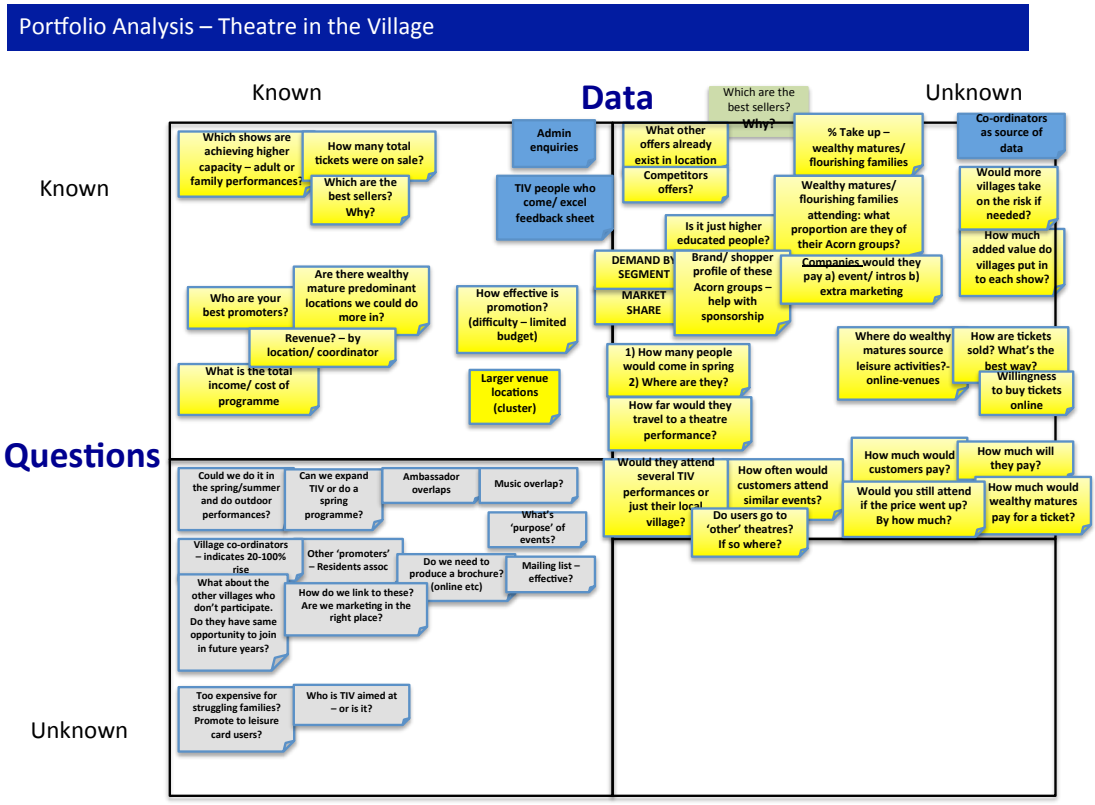


Figure 6-2 Two illustrative question analyses produced (see Appendix D.1)

Showing him the analysis for three propositions only adds fuel to the fire though. Noting how many questions and data requirements have been generated for just three propositions, he exclaims: *there are too many questions. We are not short of questions, we always ask questions, when do they actually get to make a decision!*

Andrea and I both agree there are too many questions. I explain that this was a precursor to identifying relevant questions to pursue. Based on our analysis of them, the number and variety of questions seem to point to a significant lack of clarity on the part of participants about the business model underpinning existing services as well as the new propositions being developed. It was also the first time the Leisure team had engaged in identifying customer questions in relation to their propositions, let alone which ones were most important. Nevertheless, we did identify clusters of questions and themes across propositions, particularly in relation to customer pricing and price points.

This points to opportunities to combine data gathering and analysis across propositions where similar customer groups are being targeted. The next step will be to prioritise key questions so we are a step or two away from undertaking any real analysis. The workshops also identified shortcuts to very quickly obtain relevant data and address several questions at the same time. By way of example, for Theatre-in-the-village, the Leisure team could get a group of the theatre coordinators together for a session without much effort, which hadn't been considered. For us, this underlines the current lack of customer orientation and the considerable iterative learning and coaching required for Leisure to become more proficient at coming up with good propositions and market testing them.

Albert echoes this, pointing out we are starting from a very low base in terms of commercial knowledge within teams. Separately, a business model training programme has been proposed, and while ostensibly aimed at introducing the business canvas and modelling required to develop new propositions, it also aims to get council teams to start thinking more commercially, something he reckons they are woefully poor at.

How much analysis is enough?

Restating his concern about indecision and delay, though with less force, Justin grudgingly seems to acknowledge our arguments. He still points out that what he isn't seeing in what we've shown and explained is – at what point do we decide we've got enough information to actually make a decision?

It suddenly dawns on me that this may be where a similar question from Patricia at last week's project meeting may come from. It seems there may be a growing concern at board level about the delay in launching business plans and more data gathering and analysis will further delay launches.

I acknowledge that Patricia has raised this question with us and we are working on a set of criteria and key questions to address for each proposition in readiness for launch. I mention this will probably be in a checklist form for teams to consider and is likely to consider not just market or customer questions but

also finance questions, for example around clearly understanding the cost and revenue/profitability drivers for a proposition. Andrea mentions that this is intended as a resource for the teams in the Engine Room.

Justin emphasises that he thinks this just needs top-down leadership and direction: to call it and say we have enough to know it is a runner now, it is ready to go to cabinet now, even if not all the items on the checklist are ticked off. Albert indicates this can be reviewed on an ongoing basis in the Engine Room and Andrea acknowledges that he and the CEO are the only people who can decide they are comfortable taking something forward at that stage, stressing that wouldn't be for us to call. Albert points out that doing so without ticking everything off would represent a risk and, (laughing), that's a lot of gut, let's be honest... Justin counters, somewhat defensively, *no, judgement, in terms of being balanced in the need to do due diligence and calling time, we have enough now to pitch and run with and see if it will fly. That's where the judgement comes in. Wandering down to the Engine Room, saying there's enough ticks on that, talk me through the proposition.*

I agree that the Engine Room is all about trying to reach that point *faster*, given the transparency of emerging thinking to everyone and what is still outstanding too.

Finally, that provides a neat opportunity for Albert to steer the conversation onto the safer, more productive ground of logistical planning to get the Engine Room up and running, the original focus of our meeting after all.

Soon we are into the cut and thrust of identifying and clearing an appropriate Engine Room space, planning to move teams, secure full time presence and resources to be based in the room, and how best to deal with likely pushback. Also, how to ensure we keep the Engine Room fresh and top of mind, promoting footfall through the room, coverage through internal communications, senior visibility, etc.

6.1.2 Debriefing afterwards on the way forward

Catching up in the cafe canteen afterwards, a common spot for us if it isn't too busy for a private conversation, Andrea and I reflect on the meeting over a latte for me, and a tea for Andrea. I am concerned if my irritation with Justin showed but no, according to Andrea, not at all, she thought I handled his challenges very calmly. I jest that obviously my consulting experience and training has paid off, handling lots of difficult clients over many years...

Turning more serious, we are both struck by Justin's unexpectedly aggressive, challenging tone in the meeting and wonder if his clear impatience is about the apparent lack of progress launching any new revenue generating services. Andrea is thoughtful for a moment. She is concerned that senior management underestimate the scope and challenge of trying to shift council employee mindsets to become more commercial. They seem to expect everyone to just 'get it' and then 'get on with it', straightforwardly coming up with new services that are commercially 'savvy'. There doesn't seem to be enough coaching to help staff get there from what is a standing start for them. She is genuinely sorry for them. I smile, as, for all her political savvy, she is a softie really, reflected in the causes she has campaigned for. She now reveals that even the planned business canvas training rollout is apparently in doubt, on cost grounds (they will need some external help to deliver the sessions) and given the time it will take.

We also quickly agree to alert Patricia to some of Justin's challenges, especially to the market insight approach and prioritisation, so that she can give some thought to how she wants to handle this. At least there is also good progress to report on getting the Engine Room off the ground. The challenge of how much market insight (or due diligence more generally) looms large, and we agree that I prioritise developing a straw-man framework for discussion at our next project meeting. I have already been giving this some thought, drawing on some earlier due diligence experience. We go on to discuss that going ahead with limited market testing or due diligence is really a question of risk appetite on the one hand, as well as being able to demonstrate adequate governance to

stakeholders on the other, to avoid exposure to financial and reputational losses. We agree that a solution may be to go ahead on a pilot basis, cautiously, like proceeding slowly at an amber traffic light, validating critical assumptions as you go, before scaling up, for instance. This analogy sparks my visual instincts – an idea for a traffic light coded risk-impact matrix emerges and I make a note for later. It also prompts the idea to get internal audit involved to assess proposal readiness, as a somewhat independent check.

We turn to considering the next few months. Earlier this week Andrea had dropped the bombshell that she has been offered a fantastic job as Marketing strategy lead for a charity, for a cause she is passionate about, getting back to her first love really! She is already finalising an insights workshop pack to hand over, refining our earlier pilot workshop approach into three phases. This starts with a preparatory stakeholder meeting to set expectations and agree on some preliminary Acorn analysis to focus the session on particular, selected propositions. We also stress that this is their session and insist on it being opened by someone senior on their side to outline the objectives and importance of the session. The actual session itself includes a short Acorn data briefing from Tanya but mainly focuses on some facilitated engagement with the Acorn data for the propositions selected, using the market side of the business canvas to identify key customer unknowns and questions. This was led by myself and Andrea and we also got the team to identify a 'critical friend' to invite along to represent a fresh pair of eyes. Based on our debrief, we've also added a follow-up meeting to focus on prioritising and agreeing a market insight action plan, informed by some post workshop question analysis. There just isn't enough time to address this properly in the session itself and it also doesn't really allow enough time for reflection and analysis either – we'd be shooting from the hip.

I realise and say that I will miss her! We have worked particularly well together.

We start discussing who would be good to help facilitate these sessions. Our most experienced Marketing facilitator is Helga but she is particularly busy (so hasn't been included in the project). She has also been critical of the Acorn led

approach though, instead preferring to use focus groups to gather customer data. She has in fact been working with the building services group very successfully, using this approach to test some of their new propositions. We agree there is value in using both, to complement each other. Considering other possible candidates, the only other realistic option is Alex, also with considerable Marketing experience, though less confident as a facilitator. Hmmm.... we think the best approach to recommend to Patricia would be to use Alex as the Marketing 'anchor person' in the Engine Room and leverage Helga to run the workshops if possible.

****

The train journey home is a good chance to reflect on an eventful day and make a field journal entry, punctuated by brief pauses to watch the rather picturesque English countryside roll by.

Justin's impatience and urgency surface as potentially important and I wonder again about his expectations of quick answers and insights. How realistic is this? I'm reminded of my recent meeting with OtherCouncil, at Andrea's suggestion, which seems interesting in this respect and I scan my notes from my meeting with them. They have far better resources, systems, data experience and a very focused objective of migrating to customer online self-service delivery wherever possible. This seems a more bounded challenge than GoCouncil's more creative and open-ended search for new revenue generating propositions, and also more familiar territory for participants. Nevertheless, it still took OtherCouncil a long time to get a handle on the data necessary, how best to organise them to be useful and then start using them to gain the insights they were after; they still see their data and use of them as evolving, even now, though relatively stable.

OtherCouncil want to be able to demonstrate that the service changes will not adversely impact on vulnerable citizens without access to the Internet, for instance. With this in mind they realised they needed good data about where the vulnerable citizens were, so that they could target and address access for these groups (e.g. via libraries). While Acorn data were useful, they only

focused at the household level and in some service areas they were interested in particular household members too. So, they enriched the Acorn data with some of their own, from various internal systems to provide and track service provision. It took them two years, working across various service departments to agree on relevant dimensions to measure, how to measure them, let alone how to integrate the data and allow enough flexibility for different departments using them. In the end they developed a very flexible index of deprivation, allowing different departments to weight different factors differently, depending on their relevance for their service delivery, but still using the same underlying data. Nifty and quite sophisticated I thought!

I am also amazed that this was all done mainly by the data team lead using Excel and visual basic tools, working with some data focused end-users from each department. He had avoided going to IT, given all the process and change constraints he felt they would bring, until he had something relatively stable. Now that it was, he could see value in their helping to automate some of the straightforward stuff that took loads of time to do using end-user tools, and to host the solution robustly, provide resilience and back-up arrangements.

I contrast this again in my mind to GoCouncil. From the outset I had wondered at the absence of any IT involvement, either from the IT team or an end-user with a good view and access to GoCouncil's data and some facility with end-user data tools. While Tanya was proficient with Excel, and the geographical mapping tool, she shied away from SQL type programming and didn't have a particular interest to go digging into or playing with their in-house systems data. I floated the idea as an opportunity for her but she wasn't biting. This reminds me of my brief meeting with the CIO, Dennis, recently, given my concern at the absence of any IT involvement in the market insights initiative. Dennis explained that the IT team has been completely focused on trying to agree and move to a shared services model with another nearby council to streamline IT support costs, so have been extremely stretched. They are quite a small team and don't have anyone with a data focus. They also recently lost their one reasonable business analyst cum project manager and didn't really have anyone to join the

market insight project. There are so many other competing priorities too – for instance, the website initiative, which has ground to a halt over the past few months for various reasons. Asking for a data overview, to see what customer data are being held where, the best I could get was a systems overview map on a page. However, this only shows various package applications and how they connect, with no details on the data passed or captured. Dennis does point to a recent customer data audit undertaken by Marketing that I should look at.

When I discuss the vision of becoming more customer data led and the need to build up some data capability to help realise this, he seems open to the idea, although when I talk about mining data across systems his immediate, emphatic response is no. He cites data protection and data privacy concerns, as the council has not collected the data for that purpose or with express consent. I am struck by the rush to justify a no answer, rather than engage with the valid concern, to find a solution, perhaps looking at getting opt in consent, anonymising and aggregating data, etc. This seems so at odds with my private sector, 'can-do' experience that I wonder if this cautionary principle is a hallmark of public sector thinking. Surely not...

He seems much keener on reviving the idea of implementing a CRM solution to address customer data, and I know Patricia has already expressed an interest in this too. However, the costs are considered prohibitive, especially given GoCouncil's financial challenges. My review of the recommended customer data audit, a very high level survey exercise with a limited, partial coverage, revealed a hotch-potch of data definitions and practices to capture customer data across different teams and services, with rather poor data quality noted throughout. I think they can make a lot of progress on customer data without investing in a new system. I'm also not sure they have a clear idea of what this would entail, or of their requirements, i.e. what data this would provide and how these would be used, let alone what specific benefits this would enable. To get the most out of a new system, they would need to address the data quality and coordination issues for one. Trying to address these and their most immediate data requirements with what they had, using some cheap and cheerful end-user

tools, might actually help them clarify what data and functionality they want from a new system. It may also provide a bit of a prototype to help come up with a more robust requirements list than an abstract exercise to imagine what they might need, distracted and enthused by lots of CRM feature lists. Implementing a new system seems a distraction from actually getting on with engaging with customer data already available, warts and all.

I cannot help thinking how different this data project has turned out from what I originally anticipated. It feels a million miles away from lots of sophisticated data analysis and engagement. It has been far more to do with figuring out what data are required, what data we need versus what we have, clarifying and prioritising myriad questions to address, achieving alignment across teams and even getting recognition for the need to ground new proposition assumptions in customer and market insight.

Perhaps this reflects my own starting assumptions and 'theory' – about how data projects work and about how my fieldwork is likely to generate new research insights too. I smile to myself about my research design. As I look out of the train window at the sun setting over the serene countryside, I'm reminded of a military saying 'no-battle-plan-survives-contact-with-the-enemy'…

6.1.3 Catching up with Marketing a few months later...

It is really good to see Patricia and Alex again! We are catching up in the cafeteria downstairs over a cup of coffee. Patricia is treating us!

How time has flown – it seems ages ago now that I finished my fieldwork, completing one or two more workshops with Alex, after Andrea left, so that she would feel more comfortable facilitating them, and sharing my question analysis from these with Patricia to take forward with the relevant teams.

Albert unfortunately cannot join as he is on holiday but had emailed me to say the Engine Room was going really well. They had already identified quite a lot of savings as a result of the greater engagement and focus it prompted, although he acknowledged that running it required a huge amount of energy and had been quite draining, especially initially. He had been relatively silent

about new propositions though, so this was one of my questions for Patricia today.

First though, I get an update on the team. It has been a very busy few months their end, especially after Helga left them, taking early retirement due to illness, and Patricia is still trying to find more staff. Alex is positive about the Engine Room market insight engagement, though in fairly general terms.

When I ask about new services launched, they become a bit conspiratorial, checking to see we have enough privacy. Patricia explains that the new proposed nursery service had been presented to the board and been shot down in flames by one of the elected board members. It hadn't even been about the lack of market testing, although she had not been happy with the level of competitor analysis. Rather, the business plan hadn't taken account of premises costs and had simply assumed these would be free from GoCouncil. I was immediately reminded of my due diligence questions and longstanding concern that the Finance team weren't sufficiently engaged to provide this kind of vetting and challenge.

While this crystallised a shared moment of I-told-you-so *schadenfreude* around the table, on reflection this felt somehow hollow. Certainly, on one level this vindicated Patricia's argument for the need for market insight and testing. On another, I still wasn't sure how well this was translating into data engagement, analysis and insight. When I went on to ask how useful the workshops and related question analyses had proved, I established that not a single follow-up meeting had occurred to agree and prioritise any formal market insight data collection and analysis. I couldn't help thinking this did not bode well for generating and using customer and market insights to launch successful new revenue propositions.

6.2 InfraDig, an emperor short on clothes and long on elephants...

It is a great sunny day, and Barbara has secured one of the glass-walled corner offices, with impressive views over the city, for our meeting. The meeting is to debrief on the recent workshop to map out the asset data value chain and dependencies. It has taken some time to line up with Donald and Mallory's diaries. I will miss her enthusiasm and support. She is great to work with and knows everybody in InfraDig, Yet, after six months on site a few days a week, it is also a closeout meeting, my swan song.

How quickly the time has flown! What an interesting journey it has been too, so different to what I imagined at the kick-off meeting in January. Not that I had very definite and detailed ideas about what might happen. Still, I had an implicit expectation of neatly comparing and contrasting two very different data projects in the same context: a more detailed asset data KPI project, which seemed short and bounded, with a more strategic Information Management Strategy project. The latter seemed particularly interesting in scope and ambition, given InfraDig's ambition to create a 'virtual infrastructure' to inform and improve maintenance.

However, the experience of the past few months was certainly not neat. Well, the asset data KPI project was perhaps the neater of the two projects. Casting my mind back, the KPI project seems such a long time ago now, when I was still relatively new to InfraDig.

6.2.1 Being a fly on the wall during the Asset Data KPI project

Straight into the detail

The kick-off meeting for the project was held in a similar glass walled meeting room, with a square table to seat about eight. The core asset data team were all there: Ralph, the focused and business-like team lead; Matthew, the eB superuser, given previous eB asset data project experience and an interest in database development and Excel; and project stalwart John, who while new to asset data, had been with the broader project from the start and would

coordinate the remaining asset data staff, many of whom he was still in the process of recruiting. James, their resident consultant steeped in asset data management, who had suggested my involvement was also there, as well as the high-flying Pierce, who had led the asset data induction session and was the chief technical engineer – Mallory's right hand, having advised a previous well-regarded construction data initiative. He may well have been standing in for Mallory who couldn't make the kick-off. IT was represented by Anthony, the new head of development, though this initial high level requirements phase was being led by Paul, an experienced big-four BI consultant, who had led a similar KPI project for another InfraDig area for Mallory, and had been on the wider project for over a year now. Being new to InfraDig and construction in general, I remember being relatively quiet. In fact, I was enjoying not having to lead the meeting or project – my normal consultant role in such a situation. It freed me to be an attentive observer for a change.

Ralph started by giving a brief overview of how asset data were structured. These were based on the AD4 forms, capturing the formal data requirements for contractors to provide. These were still in the process of being signed off with the operating companies, which would eventually receive the asset data. He showed Paul a related schematic on his iPad, as we couldn't get the overhead projector to link up.

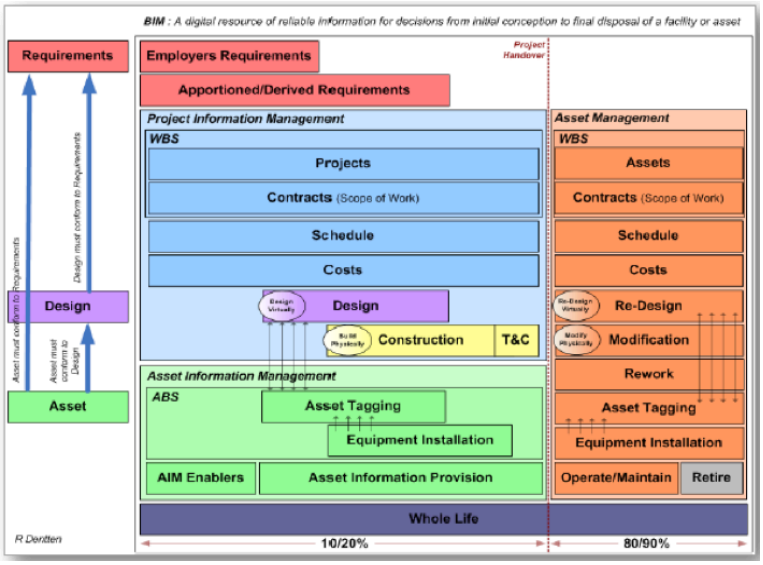


Figure 6-3 Asset Data overview reference (Source: InfraDig)

Using this as a reference point, Ralph explains how assets and components will be physically tagged, have serial numbers and barcodes assigned for tracking within eB, and how this relates to the virtual asset data structure, work breakdown structure and hierarchy. Clearly, quite a lot of thought and design has already gone into this and several clarification questions were addressed, e.g. how barcodes and serial numbers related to each other.

Before long, Paul steers the conversation towards which systems the KPI data will be drawn from. After a brief discussion, it is agreed it will be eB and Matthew, the eB guru, steps in to explain eB's overnight reporting and the related extract that is likely to be the basis of much of the intended KPI reporting. A data extract from eB is required due to eB functionality constraints to aggregate and summarise data, given a proprietary database and query language.

There is a brief discussion about scope and the key focus on reporting status of data collection. Strangely it seemed to me at the time, data quality isn't necessarily in the KPI reporting scope, per Ralph, though this still needs to be clarified. Given the focus on progress status reporting, Paul suggests that reports may need to be organised and prioritised by project or asset readiness stages, or the related project stage gates outlined by Ralph earlier.

At this point Matthew tables copies of a set of eight asset data reports he had already prepared in anticipation of the meeting, based on his previous asset data project using eB. James also contributes a list of proposed KPIs a bit later, based on his wider asset data experience, which Paul arranges copies of to circulate to the rest of us. These outline suggestions structure the rest of the meeting, with the discussion rapidly becoming quite detailed and technical in relation to particular status reports and KPI suggestions.

Questions are raised about reporting by contractor and construction location; judging progress against contract specific target start and end dates, in terms of a possible percentage completion KPI, to avoid such a KPI becoming meaningless or confusing. These lead to related questions of where contract and location data elements are captured, how, and where they reside, in terms

of systems, in order to facilitate such slicing and dicing for reporting and KPIs. Many of these questions are posed by Anthony, the development manager, and answered by a combination of Matthew and James, with Pierce chipping in when it comes to broader project and contract tracking systems, where some of these data might reside, as well as CAD systems data that may be relevant.

A suggestion is made to map James' suggested KPIs against Matthew's reports; though James doesn't think they will necessarily align perfectly, especially as some of Matthew's are more detailed breakdown reports. Anthony is also keen to see the underlying SQL database design if available, prompting Matthew to reiterate the lack of visibility, given the proprietary nature of eB's underlying data architecture and query tool.

At several points Paul has to call timeout on the level of detailed discussion on particular reports, given the focus on establishing requirements rather than detailed design, to move the conversation forward.

The discussion then moves on to the practicalities of how asset data progress updates, status and issues will be captured, via a messy sounding combination of workflow and Excel submissions it turns out. For several contract and project data elements this isn't clear, and Anthony suggests that the DW system may provide a mechanism to address the requirement to combine data from other systems as also likely, given the need to consolidate their data for wider management reporting purposes. Contract and project data also prompt a discussion about the need to clarify how these will be kept up to date and the need to establish clear, related data ownership. Pierce stresses the need to engage the project manager group in this, whose systems this would need to come from, which prompts a slightly awkward pause. When discussion resumes, it still isn't clear to me if this suggestion has really been acknowledged by the rest of the group and will be acted on.

Discussion returns to the question of data quality or validity, in relation to one of James' suggested KPIs related to tracking asset data issues. This centres on the need to establish generic and specific rules for these in relation to particular data elements. Matthew raises a related concern about where and how these

would be stored and managed within eB, given a recognition that these rules may need to evolve

Towards the end, Ralph, who has been relatively quiet during the detailed discussion, raises the question of the need for an audit trail. This prompts some ideas from James about how it has been done elsewhere and in other InfraDig areas, and I wonder out loud if there was a requirement for operators to check or sign off. That just sets James off about the fact that the operators haven't yet even managed to agree and sign off the data requirements. Ralph recognises there is a lot of complexity and related unknowns, with requirements likely to shift, so is keen to prioritise 'bread and butter' reports and park those that were likely to change.

Paul wraps up the meeting stating he feels it represents a good starting point. He also requests a follow-up session in more detail with James and Matthew to consolidate their ideas into a consistent view ahead of the next session.

Per my subsequent reflection on the meeting, it seems almost as though James and Matthew were presenting cases for their ideas to be included. Anthony and Pierce, on the other hand, were trying to add value by probing logistical issues to flesh out hidden requirements and help address some likely integration issues. Pete and I both seemed to be trying to get up to speed on the detail, while we were clearly also shepherding the process towards the project outcome that would need to be delivered at the end of this requirements and initial high level design stage.

Finalising a Functional Requirements Document

Paul has already circulated a PowerPoint high level requirements deck, consolidating James and Matthew's contributions, for review ahead of the second session. Attendance is down to the core asset data team for this and the remainder of the project meetings, though Mallory is able to dip into part of the second meeting and one of the second meetings.

Mallory makes a useful intervention early in the second session, to arrest another dive into the detail by Ralph about the need for contract details, and by

Matthew to clarify which of two different location data elements should be used. He contextualises how the KPI reporting is likely to be used in management review meetings with the wider area and project managers, so that they can address this on the contracts they are responsible for. Using the example of a particular contract, he brings it back to the practical questions that the reporting is likely to be addressing: 'how many [asset] tags does contract X need to deliver and how many have they?'

In passing, he also suggests that they will need to use contract references to address joint venture complexity, which seems to settle Ralph's earlier question. There is definitely a sense of deferring to Mallory's judgement as overall sponsor and chief engineer, though it helps that his argument has merit too. Shortly afterwards he needs to duck for another meeting, clearly a busy guy.

The discussion immediately reverts to clarifying location details, what represents completion of an asset and where these data will come from. Mark raises the question of reporting by asset class, which hasn't come up before, which prompts me to ask the question of reporting by criticality, as data about some classes of asset may be more important than others. Ralph agrees this is important, though James stresses that operators haven't yet agreed data requirements, let alone which ones are critical. Indeed, some of these themes recur over the remaining meetings.

Paul stresses several times that this stage is about capturing the requirements rather than how they will be addressed and moves the discussion onto which reports are required at which levels: area, project manager, contract, etc. He then does start discussing high-level design options and requests Matthew to provide details of eB field elements to analyse by the subsequent meeting.

Just before Mallory leaves he stresses that, overall, the InfraDig project had 33 KPIs already, which the management team thought too many, so he doesn't want the project to have too many. This prompts a discussion about trying to get to a one-page dashboard and perhaps combining KPIs. I suggest the idea of a composite index, which triggered further thoughts about Red-Amber-Green (RAG) colour coding, and more detail being provided where indicators

are Red, etc. This might also allow flexing of the KPI index over time to reflect changing priorities. Another suggestion to consider providing a more visual KPI or RAG report, using the map of the infrastructure commonly being used as a reference by the rest of the project and the location data mentioned, meets with much less enthusiasm. Instead it prompts comments about not sprinting before walking.

Following on from this, Ralph asks if a balanced scorecard approach might be relevant and should be considered, as he's heard a lot about it but doesn't really know much about it. He directs the question at James as the resident expert who doesn't seem to think it was relevant in this context. *To my chagrin, I now remember agreeing it wasn't relevant and offering to provide Ralph with more details about the approach.* The discussion is fairly dismissive and effectively parks the idea so it isn't really considered further.

As I later came to appreciate the people dimensions and dependencies involved in delivering high quality asset data, over the course of the following month or two, I think this had considerable relevance. While I did take Ralph aside later and discussed the potential usefulness of considering different measures for people versus process and outcome dimensions, as examples, the functional requirements had already been finalised. Hopefully this could still be revisited in subsequent detailed design phases.

Meanwhile the steamtrain process to produce a functional requirement specification continues apace and is wrapped up within a further meeting or two. By the following meeting Paul has already completed a detailed Excel analysis of eB datafields available, highlighting several missing attributes for the reports we are after. This throws up a question about the difference between asset functions and classes and how these concepts relate, given distinct eB datafields for each. Ralph reaches for his iPad to clarify this with reference to an example data dictionary description of an air conditioning system, which illustrates their hierarchical relationship.

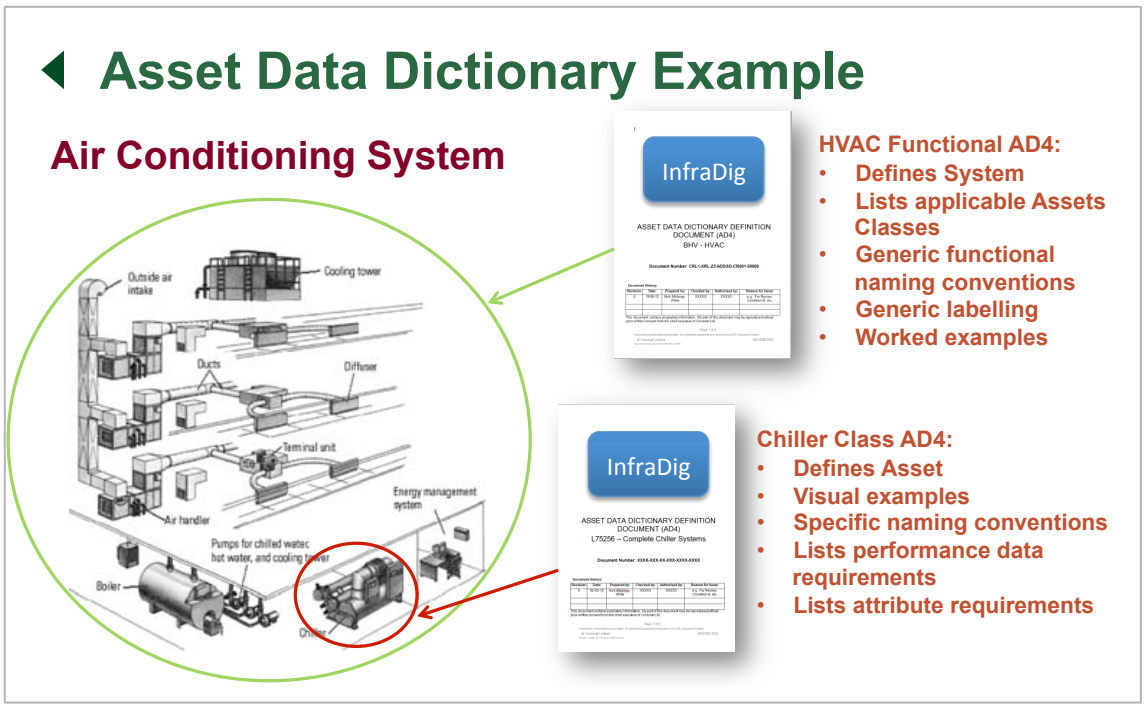


Figure 6-4 Data Dictionary example of Functions and Classes (Source: InfraDig)

Paul has also mocked up some dashboard reporting, which he's even sense-checked with one or two project managers he knows from previous InfraDig projects. It strikes me again as strange that project managers haven't been included in any of our discussions, in spite of Pierce's stressing the importance of engaging with them. This seems very much an asset data project, keeping discussion within the tent so to speak.

6.2.2 The wider IM project becomes more of a hands-on quest

The asset data KPI project produces a neat outcome, with project processes boxes ticked, related consulting job completed and billed, even if no further progress on detailed design or implementation is noted by the time of my closeout meeting several months later. By contrast the information management strategy initiative was an altogether more slippery animal.

Poised for take-off, then stalled

This wasn't apparent initially. Donald, the experienced and fairly senior manager pulled me into an opportune meeting he was having with their big four IT partner on my very first day on site. Serendipitous I thought. On entering the

meeting room I was struck by both Donald and the big four director sporting iPads. As an aside, I wonder if this is a signal of status at InfraDig and think I must remember to bring and use my iPad going forward, rather than my more pedestrian audio recording pen and accompanying notebook. We only have a short thirty minute slot, so after brief introductions, we get straight into discusssing the project scope and approach. A consultant accompanying the director has pulled together a brief, one-page roadmap approach.

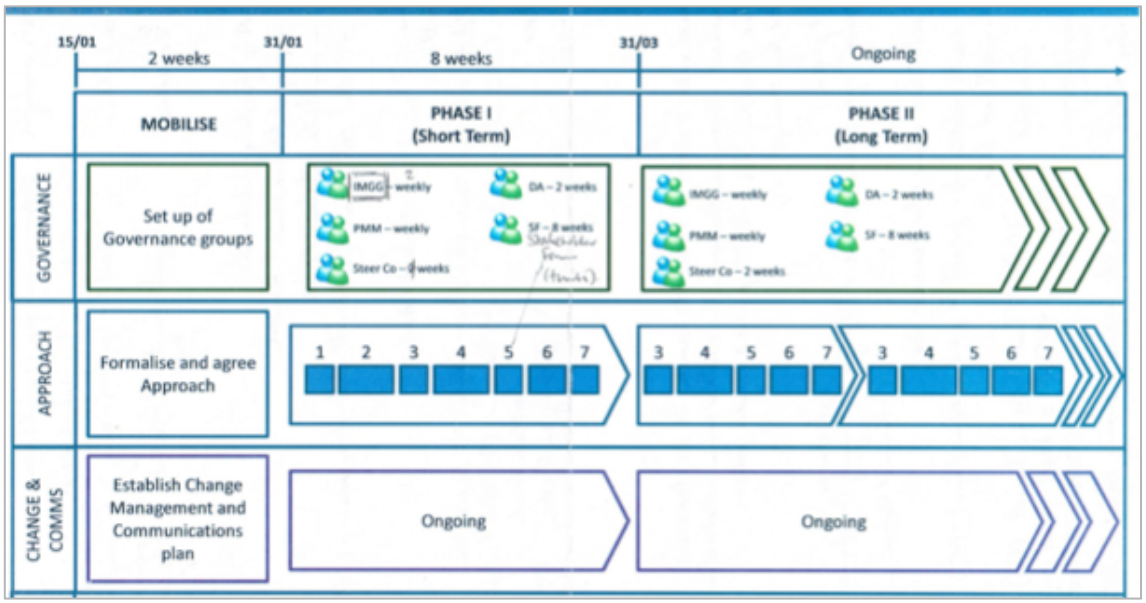


Figure 6-5 Extract from one page proposed project scope and approach (Source: InfraDig)

The director uses this as a reference during the discussion to outline a proposed seven-step approach for an initial Phase 1, followed by a longer term Phase 2. The first two steps will design some straw-man artefacts, as proposed outputs from the project, meet with some key internal and external stakeholders (one key operator was mentioned in particular), and establish a governance and related funding structure. The subsequent steps three to seven, will be repeated as cycles for particular functions (internal) and stakeholders (external) as part of Phase 2, typically eight weeks for each. These would flesh out requirements. The ownership and funding structure emerges as important to resolve in terms of the key project and organisational drivers, multiple stakeholders and related funding for the roadmap, both initially during the project phase and subsequently over the operational life-time of the data.

Donald likes the roadmap and discussion moves to what a two-to-three page business case proposal might look like. A snail shell, with a continuous, tightening circle is mentioned as a good visual metaphor to use to illustrate the gradual project clarification and wider roadmap. He needs to clear some time to sit down and pull this together but thinks he can do so tomorrow. He will then circulate a draft for us to review.

The big four director offers to free up a consultant to help if required, and they then go on to explore potential ownership and reporting lines for the initiative. While this isn't decided on, it is agreed it shouldn't be the head of IT, who reports to the financial director, but rather a more operational director. The person should be accountable for deriving ultimate value from better information management.

At this stage, the initiative seems poised to get underway shortly after the scope and commercials had been agreed and approved internally.

In search of the emperor's clothes – a clear focus, rationale and benefits

Then everything goes quiet. In fact, no draft proposal or business case was ever circulated and no sign is seen of the director. I discover from others much later on that the director has been nabbed for a bigger and more lucrative project at InfraBig, though the details are sketchy.

In the meantime, I am busy having introductory meetings with various InfraDig staff members involved in the asset data initiative, and more broadly within InfraDig, and the KPI project was underway. These meetings are aimed at understanding their view of the purpose and objective of the initiative, their role and contribution, related prior experience and a feel for the challenges they were encountering. In between, I periodically chase Donald on progress. He is apologetic and mired in reviewing and finalising the overall InfraDig project portfolio for the coming year, as part of the budget cycle. One of his responsibilities is overall project portfolio oversight for all internal projects, covering both IT and other projects. Subsequently, he is drawn into helping to

resolve ongoing operational email issues with their third party IT provider. There always seems to be something more pressing.

Nevertheless, I do manage to arrange some meetings with Donald to try and help move things forward. These tend to be interesting, if slightly rambling explorations of potential objectives, scope and approaches to adopt for the initiative. During one of these enthusiastic exploratory rambles, he sums up his thinking as follows:

> *"...obviously this is over-simplifying this significantly... I think the aspiration is that all the data is primarily our source and what is good is we have quite a simple categorisation of what comes into [InfraDig], how we use it is where all the confusion lies. I think the aspiration is this all moves into a world where we have some kind of content database, which has got content in it, which will be... which somebody will be interested in looking at... mainly [the operators]. This is what [InfraDig] would like to produce, focused on content, not documents"*

Clearly Donald sees the IM strategy project encompassing the whole organisation. It also seems to reflect and hint at frustration with their current intranet based document repository for data, outside of a few core applications for accounting, engineering and project management. The initial step-by-step approach for the IM effort to be broken down by department, completing all of them to understand their data and data use, is revisited. Perhaps establishing one or two exemplary areas first is suggested. Indeed, he often refers to IT leading by example in this area. It feels distinctly like something they ought to be doing yet at the same time there isn't a compelling reason to prioritise it.

Instead, I am struck by his extremely broad and abstract framing, i.e. categorising and making ***all*** data more accessible or searchable to everyone in the organisation. It isn't bounded or prioritised, addressing clear InfraDig priorities with specific benefits in view. I am left wondering if this is at the root of his hesitance and procrastination about pulling together a business case proposal and progressing the project. Might he rightly fear, even if he never

actually comes out and says it, that it is unlikely to come across as a compelling proposition to the project approval board. I certainly wouldn't buy it.

This challenge is perhaps further complicated by the short-term nature and time horizon for InfraDig investment. Most of the benefit is likely to flow to operators. This feeling is echoed by comments by Anthony on the development side, where he would normally do things differently if he were managing for the longer term rather than a temporary project time-scale, albeit a long one.

Let's prioritise data asset data as a starting point...

Against this, to my mind, rather woolly background, I suggest focusing on asset data as a starting point. The vision of a virtual infrastructure data asset clearly seems strategically important with a broad maintenance benefit in view, a true Emperor in terms of ambition too. Useful templates and approaches developed in this area could then be rolled out more broadly in due course. Also, Mallory seems eager and a natural sponsor.

Mallory and I bonded early on in our interview around neat iPad apps to capture meeting notes and drawing visualisations. More importantly, he certainly seems to understand the importance of aligning people around his vision of a virtual infrastructure. He also sees people as key to achieving the broader aspiration to change and improving maintenance practice using easily available asset data. In fact, he's already invested in a pilot initiative to produce a neat 3D application to demonstrate how asset data can be made more accessible for maintenance teams. He describes this as follows:

> *"So for me, when we go into the maintenance world - and one of the reasons I have that dinky little app which shows bits of wall coming off and barcodes, is because I want to get some of our sort of old-fashioned, sort of lever arch file and plan chest for drawings maintenance friends into the... to show them a view of the 21st and 22nd century, and the way which data could be used." (Mallory)*

Mallory came across as a real evangelist! His interest in the people dimension accords with my interest in understanding the different practitioner groups

involved in the initiative and how they interact to achieve the aspiration. Certainly, he seems very interested in using some of the CoP ideas I've shared with him to move things forward.

Quite rapidly, we agree with Donald to start by mapping out the internal InfraDig groups involved in asset data, and Mallory offers Barbara to help me. She is overseeing the asset data quality assurance activities across contracts and is currently running an informal community of document controllers, focused on improving interaction and problem-solving to aid asset data capture. Donald is also keen that we map out the various governance forums encountered with an eye to potentially leveraging these in due course to improve information management.

After spending so long in the doldrums, it seems we are finally on our way... off to measure up the asset data emperor to help clothe it with appropriate practitioner benefits!

Mapping the human asset data terrain

Drawing on initial interviews, several specific sessions with Matthew and John from the asset data team, Barbara and I set about mapping the groups, artefacts, key forums and governance meetings, gathering supporting documentation and terms of reference as we go. It initially involves grabbing one of the booths in the breakout areas on each floor, where most meetings are held, armed with large sheets of flipchart paper, pencils and coloured markers. It feels like a return to the fun of kindergarten art projects and we get some quizzical looks. Very quickly though, given the various groups and rich level of detail, we opt to use Prezi to map the material more flexibly. This allows for easier review and updating, with views of different levels of detail depending on the audience, and the ability to show interviewees where they fit into the bigger asset data picture.

At Mallory's suggestion we even print a huge wall-chart version and put this up in the open plan office area near the asset data quality team. We provide post-it notes and markers for engineers and other asset data staff to annotate and

enrich it further. While Barbara notes some passing interest in it and Mallory promotes wider engagement with it at one of their asset data and engineering meetings, very few additions are noted. She indicates that a slightly cynical joke made by Ralph about it hasn't really helped... It also doesn't survive the office desk move a few months later. Perhaps a bit of 'not invented here' attitude. Well, clearly you cannot win with everyone all the time.

Figure 6-6 Group mapping and open plan office work area

Nevertheless, Matthew and John find the exercise interesting, providing them a bigger picture in terms of dependencies on other groups onsite at projects, different forums, etc. Donald becomes really enthusiastic when I update him and walk him through it. He wonders out loud if it could be turned into an intranet-based resource with hyperlinks to documents referenced and profiles of some key people or boundary spanners identified. His exuberance seems to be tempered somewhat by a realisation of the enormity of the scope and effort involved just to map the asset groups, let alone doing something similar for all functional areas. I remember it giving him pause, after I'd walked him through it:

> *"...is IM too large a domain to draw up a stakeholder map – is it too much? Or my view is, let's go and find out. If it is we can break it down into Doc Control, CAD, Technical, versus non-essential information."*

Clearly, it hasn't completely dampened his enthusiasm.

It certainly gives us a much better understanding of the asset data ecosystem of groups involved, key touchpoints, forums and some pivotal people who act as coordinators, relationship managers or boundary spanners. It reveals three broad practitioner groupings involved within InfraDig itself: Design Engineers, the Asset Data Team and The Document Controllers. In addition, construction contractors are seen as external but closely related to the core engineering and project manager groups, while the maintenance and operators are also identified as closely related to the engineers, most closely via the operations delivery team. These groups are shown in Figure 6-7. These groups specialise in different aspects of creating the asset data artefact, with engineers leading the way in designing the broad design documentation structure and loading original design specification docurmentation into the CAD and eB systems, the asset data team defining and organising and managing the data in eB, tracking and ensuring appropriate data quality, while document controllers capture the related data and updated documents into eB provided by the onsite project teams.

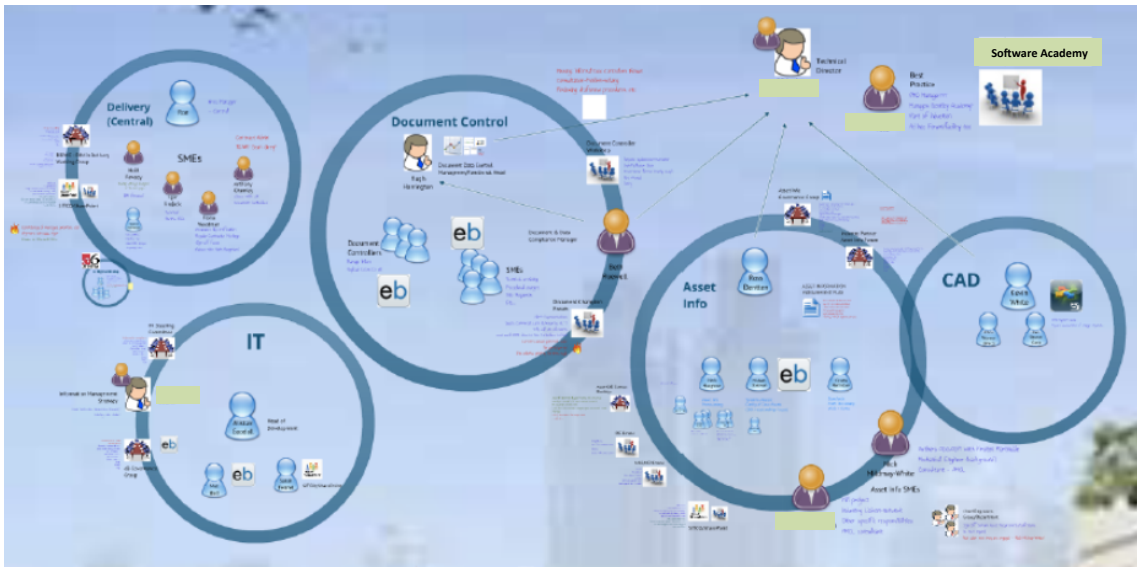


Figure 6-7 Main InfraDig groups identified in relation to the data initiative

The exercise reveals a real divide between document controllers, embedded onsite with contractors collecting asset data on the ground, and asset data, engineering and project management groups, based at the InfraDig main office. Barbara is one of the people who spans this divide given her data quality role,

which involves regular site visits, and she notes that the local document controllers refer to the main InfraDig office as the dark tower...

We encounter several surprises, bumping into various *elephants* in the course of our exploration and mapping exercise... which take on more substance and come more clearly into view during our subsequent benefit management workshop. Indeed, they are at the heart of the challenges and next steps being discussed at the closeout meeting today.

Different groups, different agendas

The first elephant we encounter is that different groups are definitely not all on the same page as Mallory, as to the purpose of the asset data collection, or the related benefits of creating a virtual infrastructure artefact. He seems to be absolutely right to be concerned about the importance and need to change and align mindsets.

On the one hand, Mallory's vision and envisaged benefits become clearer over the course of our mapping and various related discussions, further enhanced by digesting the broader industry BIM guidance. Through an introductory meeting with the IT manager I had become aware of a related government industry initiative called Building Information Management (BIM), which was becoming mandatory for public infrastructure projects. InfraDig want to demonstrate they are practice leaders in this area. Strangely, neither Mallory nor Donald had mentioned this to me as a key objective in relation to the asset data initiative. In fact, none of the asset data team members seem to be aware of BIM either and I find myself explaining it and sharing the related guidance with them.

From this two benefit rationales emerge. One is to realise benefits from the data InfraDig collects through informing 'smarter construction' practices and approaches, led by a combination of engineers, project managers and construction partners working together. To this end, a number of innovation forums and initiatives are in place or underway, with nominated champions and coordinators in various teams. Some of these initiatives include making data electronically available to construction teams via portable devices on site,

facilitating updating and annotating design drawings ('red lining') straight onto electronic versions of the plan, making it easier to capture and submit changes quickly or even in real time, enhancing timely transparency for other impacted construction partners and the designers.

However, given that much of the design planning is already complete (and essentially codified in the project plans and related contracts), the main focus or emphasis at this stage is now on construction coordination and execution, rather than redesign or optimising or improving design on-the-fly (i.e. on efficiency rather than on generating new insights). Overall construction design learnings and improvements are likely to need to wait until the next project to be implemented and tested.

The second rationale for realising value from the data is informing 'smarter maintenance' practices to enhance maintenance cost-effectiveness and reduce overall lifetime cost of ownership. Therefore, this is the rationale we focus on in our exploration. This requires the infrastructure or train operating companies (rather than the construction project or partners) to use these data to enhance their practice. While the asset data team seem to be broadly on the same page about this objective, more operational contacts we interview are far less clear about Mallory's vision.

The operations elephant...

Most notably, the operations team, responsible for managing the physical handover of the infrastructure to operators and the related handover plan, have a very different view of its ultimate purpose. Tom, the handover plan manager, in response to a question to clarify his view of Mallory's asset data vision and virtual asset data artefact, responds as follows:

> *"I know they mention the term digital and the real [infrastructure] and distinguish between them... I'm not sure I entirely understand... what it is that this is trying to do and why it is important. I mean I know at a nitty gritty level... to make it easily accessible to our approving body without them*

having to spend years doing it, and being a legacy system for operators to see how we assured it..." (Tom, operations team handover plan coordinator)

For operations it seems much more about facilitiating quality assurance to ensure readiness prior to handing over the infrastructure to operators. It also means being able to demonstrate this in terms of audit trails, should any subsequent contract disputes arise between parties, when having the relevant evidence easily to hand would be useful. I gather this stems from bitter prior experience on similar earlier projects.

The operations team represents an interesting innovation adopted to promote knowledge transfer to operators and to ensure an operator perspective is readily available to the project. It therefore mainly comprises secondees from the various operators who will transfer back to operators on completion. This contrasts with the rest of InfraDig, which is staffed largely by contractors who have been involved in similar large construction projects.

In addition to the physical infrastructure, the team are also responsible for the operational handover of the virtual asset data artefact and this is reflected on their handover plan. However, no asset data staff have been seconded from operators onto their team. This strikes me as strange given likely data integration issues. This might well ease some of the data design requirements signoff challenges and delays rumbling on in the background, never mind the usefulness of knowledge transfer in this area too.

It alerts me to the possibility that asset data might not be perceived as a high priority by operations or operators. Certainly, based on the interview with Tom, I am not sure how much visibility and traction Mallory has achieved with the team. The operations elephant seems to be dressing the emperor in very different, rather ordinary clothes, and some essential items of clothing seem to be completely missing too.

The challenges of herding several elephants...

The asset data KPI project has already alerted me to the other large elephant or engine of activity within InfraDig, the project delivery organisation, supported by

various contractors, perhaps baby elephants in their own right. The asset data initiative seems to be hitching a ride on this particular elephant. However, based on our interviews it is clear that this elephant is overwhelmingly focused on delivering the *physical* infrastructure. While the idea of building a virtual data asset equivalent seems a neat new idea, it is also a distinctly secondary consideration to the main and very tangible objective of delivering infrastructure that works.

We establish that project managers and their project field engineers have considerable discretion around contractual stage gate acceptance, as well as the format and quality of related as-built design documentation and asset data. Asset data also represents only one among many potential snaglist items or issues to resolve at completion and it isn't clear that the quality and format of asset data ranks as a high priority in related negotiation. For instance, some of the infrastructure already delivered by one contractor has simply been provided in hard copy lever arch files, the traditional approach, and this has been accepted. The asset data team are now pragmatically considering digitising this material and loading it into eB themselves to remedy this.

BIM and creating a digital asset data artefact clearly reflect a significant change to traditional approaches to construction. The engineering team, which includes the asset data team, had sought to beef up contracts in this area, though many had already been agreed early on. Certainly, they had pushed with some success for eB to be specified as the delivery mechanism for such asset data, as well as related design and operations manual documentation. The CEO's support helped here and for funding the asset data initiative. It is also likely to be important in due course during management review to ensure attention is paid to the proposed asset data KPI reporting. It seems it is personally important to the CEO that InfraDig is seen to be leading the way in terms of the government's BIM initiative. After all InfraDig is a high profile project with considerable public funding.

Engineering and asset data have also pragmatically worked with existing practice and documentation standards to make it easy to provide the asset data

they are after. These strike me as *institutionalised* boundary documents, which we capture in our mapping exercise. They could after all specify XML tagging of key data elements within documents for easier extraction and later use. While introducing such innovation later isn't exactly precluded, this will require clarification with operators as to particular data elements of value, establishing coding standards, etc. None of this is in place or even under discussion.

Contractual and other organisational barriers and tensions everywhere

All these elephants pursuing their own ends surface, in several conversations, as complaints about silo behaviour within InfraDig teams. On one level this seems to prompt a requirement for more and improved communication between different teams. Indeed, this is one of the prime reasons Mallory had been interested in CoP that could span organisational units, and supported the mapping exercise.

On another level though, it strikes me that there are some fundamental misalignments of interest exacerbating such silo behaviour. For contractors, for instance, there seems to be a perverse incentive or conflict of interest in connection with providing high quality asset data. They are likely to be bidding for subsequent maintenance work, where their unique (tacit) experience and knowledge may well be a differentiator. The project is now requesting them to codify and transfer considerable knowledge for the operator, which might well undermine this advantage.

Similar tensions are also noted between InfraDig and the ultimate Operators around the question of minimising project construction costs at the expense of longer-term maintenance savings. This is crystallised for me by a comment during a discussion with Mallory:

> *"Our finance team have only one thing in mind and that's cost for [InfraDig] and they're not into... there's a bit of tension between our sponsors and [InfraDig] with finance... because our sponsors want us to optimise whole life costs and [InfraDig] want us to minimise project cost... that is going to remain, it won't go away... it is just accepted." [Mallory, Chief Engineer]*

Overcoming such conflicts seems to rely heavily on the professionalism of the various parties, teams and individuals involved, as well as general transparency and the involvement and vigilance of the operations team. This highlights the important role played by a handful of boundary spanners we identify in our mapping exercise. Many, like Mallory, have both engineering and maintenance experience and are in consulting or relationship management roles, so can see and articulate the benefits of asset data use to improve maintenance in terms of real examples. Interviews also reveal their personal commitment to improving maintenance as being important, and several come across as passionate evangelists. One potential problem I foresee though, is that they are all based in engineering or asset data and might not be seen as local to the communities and groups they are 'preaching' to.

6.2.3 Back to the benefit workshop and related debrief meeting...

Addressing missing benefits and related operator elephants...

These operator tensions even extend to efforts to help them derive maintenance benefit from the virtual asset data that will be provided. Indeed, I recall this giving rise to a rare and uncharacteristically testy exchange with Mallory, during our planning discussion for the benefit workshop we are debriefing today.

I was emphasising the need for more engagement with the operator asset data and maintenance teams to focus on how asset data might be used to realise related benefits, perhaps using pilot data to experiment with early, during handover, to promote learning and understanding. I thought and argued that this might even help refine data requirements, finalisation and signoff of which still seemed to be rumbling on. Well this clearly touched a raw nerve:

> *"... quite frankly, I don't want to sit in long protracted meetings… to get them all on the same page, I'm just going to tell them what they need to manage and build [an infrastructure], because our role as a contractor is not to bring the whole world with us…*

"…I just want to get to the end game which says this is what we're going to do, and say here you go, this is what you are going to get and it's going to be amazing...

"Because we don't have... time to... I appreciate there's a degree of persuasion and bringing along and collaboration that's needed... I guess I've used up my collaboration genes... particularly in this world as it is not ours to define... and yet they are up to their neck in alligators so they have no idea what the best way to manage the swamp is... so we can help them create the vision for a future world and there could be a degree of sharing and caring and doing that together but actually, the fastest way to do it is for us to, I think, propose and bringing them along and we could get into workshops together but actually... it ain't going to happen like that, I just can't see it happening like that and I actually think that... if we were going to go down that area [route] it is quite a big area of activity…"

This clearly reflects considerable frustration working with the various operator stakeholders to achieve agreement on data requirements, especially given considerable InfraDig focus on meeting defined project timeframes and outcomes. Evidently, this seems to be engendering a real feeling of time and budget pressure, acting as a constraint on more exploratory engagement in favour of arriving at a pragmatic, good-enough answer quickly.

The project timeline and associated budget concerns about more engagement effort remind me of a meeting I had overheard shortly after returning from holiday a few weeks ago in the breakout area. One of the innovation groups was meeting to discuss rationalising such initiatives to concentrate on the most promising ones that would provide the most value in the short term. My antennae immediately pricked up and instinctively I sensed that something had changed while I'd been away. I quizzed Barbara about it and a week or so later it duly emerged that InfraDig had just reported being about two or three weeks behind schedule. While it doesn't sound like much on such a long project, there is an almost direct link between time and costs on such projects so no wonder

anything discretionary is being challenged. The message had clearly gone out to concentrate on the knitting…

Nevertheless, I continued to argue for the need to focus on and clarify how benefits are likely to be realised. This might enable us to identify key dependencies, accountabilities and promote the chances of achieving these. This thinking is grounded in my exposure to benefits management thinking during my project risk consulting, enhanced by my related interactions with Cranfield, who had developed various neat approaches and techniques in this area.

COBie questions as a potential starting point for benefits

Given the acknowledged absence of clear asset data requirements and benefit objectives, the BIM guidance seemed to me to present a useful potential starting point. It includes a reference to a COBie (Construction Operations Building information exchange) framework for sharing construction data with operators. This had apparently been developed by the public sector in the US and aligns the operator data in order to address typical questions. These questions centre on design performance, maintenance scheduling, operational costs and designed versus actual energy use and costs. They reflect government objectives to minimise maintenance costs and energy footprints for public infrastructure, an emphasis on the latter reflecting a new focus by government to reduce the country's carbon footprint. Such questions seem a potentially useful way to prioritise and possibly even organise data elements, offering a useful reference point for data requirements discussion. It might even provide a basis for some pilot data exchange and experimentation. However, no-one within the asset team had even mentioned COBie or the questions and I am reminded of the KPI project discussion and lack of clarity about data element criticality, and the need to focus on these elements for data quality and reporting.

Where was the Finance elephant in all this?

Perhaps it is my previous accounting training, but COBie's clear emphasis on cost-effectiveness prompts a concern about the absence of any finance

engagement, representation or even reference by Mallory or other asset data interviewees. Surely they might represent excellent allies and a point of leverage to persuade operator maintenance and asset data teams to give this adequate attention, if that was the sticking point. I am also put in mind of James mentioning a BIM audit of InfraBig, a much larger national related infrastructure organisation. Their finance and asset management functions had independently developed completely different hierarchies around assets and costs, a duplication of effort certainly, though the worst was that they weren't compatible or aligned. Similarly, there has been no mention of energy related data elements or representation by asset data teams so far.

This missing finance element had prompted me to ask some questions of interviewees but only one, the engineering group operator relationship manager, mentions the existence of a detailed maintenance cost model:

> *"**...**in fact [it is] too detailed as it doesn't let you make a quick decision as to which way to go – people, etc... [One operator] wanted to validate it against how they do it. Interestingly, they do it per square foot... Pound per square foot."*

When I ask who had developed it, he mentions someone from James' asset data consultancy:

> *"... she is running it – trying to use [it for option selection... Currently, there are a lot of assumptions in there, which need to confirmed and validated by the maintainers because we have assumed one maintainer... their union rules may say you need two people working on that."*

We briefly discuss that this might be an excellent prompt and support for the idea of some asset data piloting and experimentation with the operators. It sounds like InfraDig had some good building blocks to use for this. As an aside, he also mentions some usage movement modelling he came across earlier that very morning. It isn't joined up though and there is little awareness of the activity more widely. This analysis and modelling also seems to be using many

assumptions. More evidence of silo activity and external stakeholder dependence too I think, something that I am hoping to surface and address through the benefits workshop.

Visioning activity preferred to COBie operator engagement

Suggesting using COBie as a possible starting point to get workshop participants thinking about benefits, however, triggers another uncharacteristically dismissive and emphatic response from Mallory:

> *"…personally, I think COBie is irrelevant to [InfraDig]… …I'm not so immediately interested, it is academic, it is of purely academic interest to me because we are not using COBie... bluntly, we are using our own spread[sheet]... …the COBie question, people talk about we're going to hand over our information, it will be in a database not in a spreadsheet."*

Obviously, I've touched another raw nerve, though one I clearly don't understand. He seems to be concerned about the spreadsheet mechanism and format rather than the idea of underlying questions. Indeed he goes on to explain his antipathy:

> *"…the problem with the use of the word COBie... the amount of baggage it comes with is seriously enormous... because the elephant in the room with COBie is that government is trying to mandate the use of COBie spreadsheets because they want a particular type of spreadsheet format to be used by all public contracts and as long as I have any breath in my body it ain't going to happen like that.*
>
> *"So if you were to talk about Asset Data requirements then that is a much better label... My worry with COBie is that... it won't get ditched... They WILL change some of the labelling and the numbers, it will evolve... Like all Germans aren't bad, there is a huge amount in there of value but the use of the word COBie is just going to get in the way..."*

Instead, Mallory is far more interested in other initiatives to get people thinking differently:

> *"…of course they cannot tell us because it is all over the place, so we do have to try and put some sanity and some rules into that world... I was talking today about wanting to create a 2025 [infrastructure vision], set out or draw up what we think a 2025 maintenance regime will be, so we can understand what those activities are because until you do that you can't really work backwards...*
>
> *"…one of the things we are doing is... I mean separately to this... we are into producing the Hogwart times – we are creating the O&M operational maintenance manuals, which we want to be rich and interesting and joining up bits of data… …we've been talking about this for a couple of years now... we haven't nailed it... [Pierce] and I [have] actually decided to become digital chippies and become digital bookcases, so we can get people to get their heads around a couple of proof of concepts…"*

He also feels they could address most of the COBie questions right now. This seems to me a strange paradox – why then is it proving so difficult to get agreement with the operators around the data requirements. Mallory is much more receptive to the idea of focusing on practical business challenges faced by particular practitioner communities and improving communication between practitioners as a starting point for the workshop:

> *"[I'm] more interested in the Communities of Practice as in there is some of the subtlety or direction for next steps... Communities of Practice have already taught us a few things and there is more in there...*
>
> *"…what are the business problems and I think what you've got in terms of the Communities of Practice stuff, there's a lot more richness to be gained about how the organisation is not communicating very well..."*

I am still left with a feeling and related concern that they are intent on coming up with *their* ideas of what clothes the emperor should wear, setting the maintenance fashions for 2025 and then selling or evangelising it to operators.

Back to the benefit workshop debrief

Everyone is now here and we are ready to go. In addition to Mallory and Donald, we had also invited Ralph as head of the asset data team. We thought we would keep it small though, to discuss internally how best to take things forward in terms of next steps, so hadn't included Tom from the Operations team for now. Tom had been part of the workshop though and had brought a useful and different perspective. Donald had sent someone instead as he was still helping firefight the ongoing email issues. Mallory had dipped in briefly but it was mainly the asset data team and James, the asset data consultant in the session.

The workshop session had been structured in two broad parts; the first focused on exploring the key benefits, identifying which practitioner groups would realise these. This had used the CoP mapping exercise, which we had pinned up on one wall for easy reference. We then had everyone do a quick post-it exercise. This flushed out their top critical dependencies or critical success factors to realise the benefits identified, getting them to put these up on a wall under broad People, Process and IT/Data headings, before reviewing and clarifying these during a group review. The second part of the session, after a short break, then prioritised the groups identified, and assessed their benefit readiness and maturity using a simple RAG scoring. We also assessed the level of InfraDig engagement with each of these using a similar scoring, before exploring what forums and mechansims existed to improve engagement with priority areas. It was an ambitious agenda for one morning but represented a good trade-off with participant availability and attention, to establish a rough starting point.

After asking after Mallory's health (he had been off ill for a few days) and establishing that he was feeling fully recovered, I kick off the meeting, using Prezi on the room wallscreen as a reference point. This outlines in overview the structure for the session: touching on the workshop structure, approach and objectives, as a reminder; then a review of the benefits value chain analysis that had emerged; and a review of the related stakeholder analysis; before finishing

up with some proposed next steps. Everyone is by now used to my use of Prezi. I am still revelling in its neat ability to put everything that might be relevant on one page, allowing us then to zoom in and step through the material at whatever level of detail the audience prefers and time allows.

Figure 6-8 Prezi workshop debrief presentation overview

A question from Mallory prompts a quick recap from me on value chain analysis and related dependency maps, to make sure everyone is on the same page about these. In particular, I stress that we only considered operator benefits rather than project construction benefits. I then zoom in to the main post-it note mapping workshop output.

Figure 6-9 Zoom-in to value chain dependency map output

I'd recaptured the post-its from a photo taken at the time, as a clearer and more flexible format for InfraDig to re-use and refine going forward. Starting with the benefits and related practitioner groups, we zoom in further on this area and stepped through some of them at a high level, highlighting the key things that had come up in the session.

More specific benefits and related groups clarified

I explained that two broad areas of benefits from asset data had emerged as important. The first relates to operating and maintenance costs (i.e. total cost of ownership), the second is around assurance that the infrastructure was ready and working. Interestingly, the specific BIM term, or reference to BIM, hadn't come up at all in the session. Finance, had emerged as expected as an important stakeholder grouping to realise these benefits. Ralph comes in here to emphasise and explain that Finance and Asset Data teams are poles apart. Malcolm acknowledges that he 'gets this', which represents a shift from our earlier workshop planning discussion. I go on to explain that the session also identified the practitioners involved in realising the maintenance benefits more clearly and specifically. The main users of the improved asset data will be the maintenance planners, which could involve considerable data analysis. Maintenance crews might also benefit from having better data available when actually undertaking the maintenance itself, though this is more likely be in relation to the particular asset being maintained, i.e. a different data footprint and perhaps less analysis.

A question from Donald prompts a short discussion about whether or not these benefits cover efficiency and effectiveness, and I suggest they certainly could cover both. I raise more of a concern about the extent to which the objectives and related benefit measures are 'SMART'. When I get some blank looks, I realise this is a consulting term they are not familiar with, so I explain that they ideally need to be: Specific, Measurable, Action oriented, Realistic and Timebound.

They agree that while some of them are defined many aren't clear and that would be something for the operators to look at. In relation to this, Ralph comes in to explain:

> *"…we also discussed the idea of making [asset] data great, is not just our responsibility. We are just data asset management, the guys we are handing over to – infrastructure managers. They should be involved during the project to ensure their systems are in place."*

Mallory responds, saying he has just been talking to the maintenance guys yesterday:

> *"…one of the things we were talking about was whole life costs, OEM manuals and things like that and ensuring we are getting and how we are getting/do deliver the optimisation of whole life cost requirements we have in the project... with the [Operator A] people... so we are doing it, or the intention of doing it, maybe the way we are going to do it needs some adjustment but there is the intention to do it..."*

Critical dependency elements and the transition handover plan

At this point I rezoom the Prezi to focus on the critical dependency elements that were identified. Here some participant surprises emerged as significant dependencies or requirements to ensure success. In particular, this included: operator system readiness and integration effort on the IT side; as well as a variety of people related knowledge challenges, related to knowledge handover to operator staff, especially more tacit aspects and experience.

Ralph comes in here to support the importance of the people aspects:

> *"…there isn't anyone looking at this and someone will need to be responsible for this at handover. He guesses it would be the [InfraDig] Ops guys but it would be nice to know..."*

There is general agreement about the importance of the early engagement of operators. At this point I mention the operations team handover plan, as a mechanism to help achieve this, highlighting that this currently focuses on more

tangible artefacts being handed over, rather than ensuring more tacit knowledge transfer. I say the benefit dependency mapping may be a good artefact to use to engage the Ops handover team to discuss these aspects. Per Barbara, Tom seemed to get a lot out of the workshop session and may well be able to help broker this as owner of the transition plan.

Mallory comes in at this point with a surprise:

> *"I think the person who got most out of it was [Ralph], who went in quite cynical and came out a 'changed man'..."*

General laughter ensues. Certainly Ralph had seemed cynical of our practice mapping exercise per Barbara's grapevine feedback, though in today's meeting he has participated enthusiastically and is being very supportive of the need to do more. This is a very positive shift indeed.

Ralph asks me to send a copy of the handover plan, as he has not seen it, which I agreed to do. This prompts a question from Donald if it is in the eB document repository. Barbara responds that she was sure it is, as Tom is generally quite good about that.

Donald also comes in to contribute that a single IT system would make some of the people issues much easier, to general joshing from Mallory. There seems to be just a bit of an edge of truth to the jesting though… hinting at longstanding tensions between IT and engineering over delivery.

Stakeholder focus areas and next steps wrap-up

Stressing that there is still a fair bit of work to do to refine and finalise the benefit dependency map, I move the Prezi and discussion on to the stakeholder assessment output. This underlines the operator engagement gaps, particularly with one or two stakeholders and with finance groups across all of them. Donald comments that we may have been generous with our RAG assessment for the IT function engagement and this may also require more attention.

The assessments generate a fair bit of clarification discussion about the relative significance of different stakeholder groups in terms of the volume of asset data

they will take on versus how problematic they may be in terms of readiness. For instance, Operator B may not represent the most asset data but certainly represents the most system and process complexity, as well as related resistance to making changes.

As the session draws to a close given the hour allocated for it, I move to some proposed next steps. This summarises much of what we've already discussed.

Figure 6-10 Zoom-in to recommended Next Steps

After a brief pause to take it all in, Mallory and Donald shift to a discussion about who best to engage with to try and take this forward, mentioning names within InfraDig and key Operators. These reflect various personal assessments and also surfaces that for Operator A, the challenge is that the organisation is still being formed, so several key positions we would ideally engage with simply aren't in place yet.

At this point, we run out of time and need to leave the room as the next meeting is hovering outside. Donald takes me aside outside to thank me for all the effort and I head off for a quick farewell lunch with Barbara to thank her in turn.

6.2.4 Epilogue...

It is with mixed feelings that I was leaving such a rich setting. However, there seemed to be several hurdles in the way of any operator engagement in the short term, which seemed like the next logical progression for further fruitful case understanding and analysis. There are also more pragmatic financial considerations, having taken a break from consulting to undertake my research fieldwork and data analysis. So per discussion with Joe, my supervisor, the debrief meeting seemed like a logical point to withdraw from the field.

Some interesting discoveries wandering off the path...

Nevertheless, it has been a particularly rewarding immersion in a rich and complex setting. In addition to the ostensible focus on the two data projects, it has afforded a wider appreciation of industry-level BIM data developments, as well as other data related conversations with various InfraDig staff, including their database administrator about the more technical challenges they were facing.

Two such conversations stand out for me after reflecting on leaving the field. The first was with Anthony about another data challenge they were seeking to address. This related to the construction project teams' collection of a vast amount of movement sensor data to monitor any potential impact of tunnelling on buildings in the vicinity. I had immediately been interested, given the huge volumes of data involved and their proposed use of QlikView, a big data tool, to help address and present the related data to the project team. While the project didn't crystallise while I was on site, our initial exploratory conversation was really insightful. It became clear that while there were considerable challenges related to storing such a vast amount of data, this aspect seemed manageable as the data wasn't required for very long.

When we dug into the processing of it, this too emerged as being surprisingly straightforward. The project staff simply wanted to be alerted in real-time to any movement noted outside very tightly defined tolerances. While this represented more of a communication challenge, this could be relatively easily addressed

using mobile devices. The data analysis seemed extremely straightforward, essentially a variance analysis.

A second conversation that had been arresting was with an infrastructure modeller based within the InfraDig operations handover team. He was modelling likely movement through the infrastructure based on designs, and evaluating the impact on the flow of proposed designs or as-built changes. This effort was highly reliant on modelling software from an international specialist provider, who had developed a complex underlying theory and model over many years. This was of course proprietary and was constantly being refined with regular updates, reflecting tweaks to this underlying model based on feedback from users in the field.

I remember these conversations being striking – highlighting how not all big data applications were necessarily analytically challenging and how different data projects might pose very different challenges. In fact some big data projects might not be particularly challenging at all, while other, more traditional, data projects might pose far more challenges. It also alerted me to the different levels of complexity inherent in different data tools being used.

A year or so later...

My decision to withdraw from the field was vindicated when I caught up with Mallory and Barbara a year or so later to see how they were getting on. Engagement with the operators was still proving a challenge with Operator A still struggling to appoint and retain appropriate staff to certain data and IT related positions, while Operator B was still proving resistant to change. Barbara though now had an iPad and had been promoted! Asset data collection was also now well underway and her role coordinating data quality across document controllers was keeping her busy and was clearly valued. Nevertheless, I remained concerned about the likely benefits that would be realised by operators from all that assiduously collected data…

7 MAKING SENSE OF MY CASES – AN EXPLANATORY FRAMEWORK

Before presenting a theoretical discussion in Chapter 8, this chapter seeks to make sense of the cases described in Chapter 6 by comparing and contrasting them. The provisional explanatory framework used to structure the comparison emerged from my initial data analysis. The case comparison serves both to explain the cases in more theoretical or abstract terms and, at the same time, serves to illustrate the framework.

The case comparison brings particular features of the cases into clearer focus, generating further insights and refinements to the framework. This section substitutes for the more typical findings section used when presenting research adopting other methods, with particular findings supported by particular data collected. Instead this section presents my sensemaking and reflexive observations based on the thick descriptions already presented and a synthesis of the data analysis described in Chapter 5 and evidenced in Appendix D.

As part of the cross-case analysis, my own research is introduced as a third case for reflexive comparison. As already highlighted earlier, my own research also pursues insight from the data I collected. Given that my research represents an individual research effort rather than a group effort, it provides a useful opportunity for triangulation. Including my case in the comparison in this way also makes explicit my reflexivity about my own practice, addressing one of the key validity requirements for ethnographic research (Schultze: 2000). Reflexive sections and related observations are presented in italics to highlight them and to differentiate them from observations and reflections related to the cases studied.

The chapter starts by briefly introducing the provisional explanatory framework used before presenting the case comparison itself. The chapter finishes by presenting a summary of the case comparison and the refined explanatory framework that emerges. This refined explanatory framework represents the starting point for the discussion presented in Chapter 8.

The case comparison starts by considering the purpose and framing of the data initiatives in terms of their underlying logics. This challenges the dominant logic outlined in earlier chapters:

More Data + **Better** Tools = **More** Insight

The case comparison goes on to examine the various ways in which practitioner groups involved shape their data initiatives, drawing on several CoP concepts, before focusing on how iterations of data engagement result in increasing clarity and learning. The final sections of the case comparison focus more closely on the role of data, tools and artefacts used in such engagement and learning, noting several barriers to data engagement.

7.1 A provisional explanatory framework for pursuing insights from data

Figure 7-1 describes the key elements that emerged from my data analysis, representing important features of the data initiatives observed. The framework provides a clearer view of the project level phenomenon – how participants in data initiatives actually pursue insight from data – than that presented earlier, in section 2.7.4, with which I approached the field (see Figure 2-8).

This also represents a good juncture to draw attention again to the fact that the terms project and initiative are used fairly interchangeably in this thesis. This recognises that not all initiatives may be defined as formal projects. The importance of using these terms though is to draw attention to the collective group level phenomenon, i.e. distinguishing it from individual level and organisational level phenomena. The clearer view of the phenomenon that emerged from initial data analysis is briefly described in terms of the provisional explanatory framework presented.

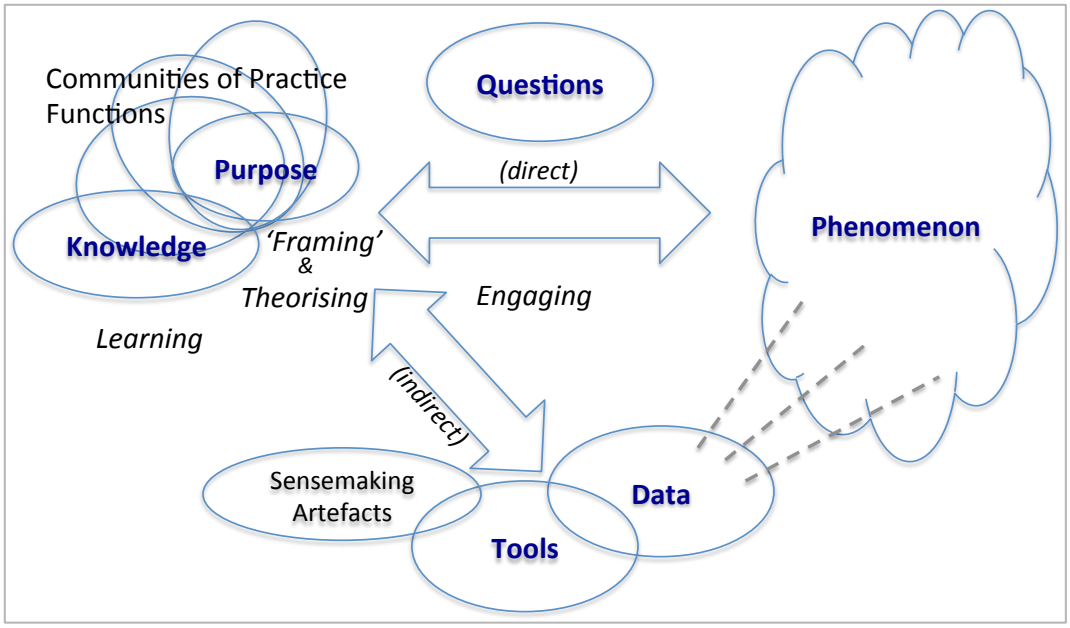


Figure 7-1 Provisional explanatory framework for pursuing insight from data

Practitioner Engagement with a Phenomenon of interest emerges as central to data initiatives. Engagement involves interacting *directly* with the Phenomenon (or practising it), or *indirectly* Learning about it through Engaging with related Data about it, or both. During such Engagement, Questions are posed about the Phenomenon (implicitly or explicitly), which are 'framed' by an overarching Purpose and prior Knowledge and experience. These are situated within a practice and organisational function context. Learning occurs through this engagement and new Knowledge is generated (or existing Knowledge is confirmed or called into question). This Learning can be about the Phenomenon of interest itself (the main objective) but it can also be about the related Data purporting to represent it, and about the Tools used to Engage with related Data.

Tools are used to enable Data collection and organisation as well as Practitioner Engagement with the Data to perform analysis, and present related findings and generate outputs (e.g. reports, screen displays, visualisations, Excel extracts, etc.). These outputs represent 'Sensemaking Artefacts' for other Practitioners when presented to them or selected and used.

These explanatory framework elements and their interaction are explored and illustrated more fully below (Capitalising them for clarity within the text). The case comparison also leads to further refinement of the framework and new elements that emerge. These are highlighted and incorporated as framework refinements as they arise. This culminates in a refined explanatory framework presented at the end of the chapter.

7.2 Clarity of Purpose – new knowledge about a Phenomenon

Very early during the analysis and reflection, I noticed that both projects essentially sought new knowledge about particular Phenomena of interest to fulfil wider Purposes. These Purposes were specific to their particular organisation, function and Practitioner Group context.

In the case of GoCouncil, the ostensible Phenomenon for which Acorn Data Engagement was proposed, is customers for new or 'revamped' revenue generating services, in particular their service preference and optimal price points for services. The ultimate Purpose though was to come up with viable and attractive business models for new or revamped services to implement, which could generate additional revenue or provide existing services more cost-effectively, achieving savings.

In InfraDig's case, one Phenomenon was maintenance practice and the main wider Purpose to reduce maintenance costs or total asset lifetime costs, through improving maintenance practice. However, other wider Purposes (and Phenomena) emerged for the data initiative, aligned to different Practitioner Groups, as well as internal and external organisational entities (e.g. the quality assurance audit trail requirement for the operations team).

InfraDig also reveals how the lack of a clear and compelling Purpose, can lead to procrastination and a lack of progress in data initiatives, as in the case of their wider Information Management strategy. The framing for this initiative is extremely broad and abstract, i.e. making all Data more accessible to everyone in the organisation, rather than bounded or prioritised by clear Practitioner

Group contexts with specific benefits to particular Practitioner Groups. This makes creating a clear and compelling business case very challenging.

In the case of GoCouncil the market insight inquiry about customers highlighted the need to also consider the wider market for the services proposed, e.g. likely competitors (and suppliers). For instance, existing competitor price points, scale and profitability, as well as the overall market size, profitability and scale requirements are also relevant.

7.3 Inquiry and Benefit realisation logics left implicit

The search for new knowledge or insight described above can be thought of as realising or enacting particular inquiry logics. My own research highlighted the role and benefit of articulating and pursuing a clear inquiry logic, captured in my research design. It identified and kept defined research Questions and the Phenomenon in view. It also clarified how any Data collected might be analysed to shed light on the Phenomenon by addressing the Questions posed.

Turning to the cases, I noted that their inquiry logics or theories remain largely implicit. For example, in the case of GoCouncil, an important implicit assumption, about which Acorn household groups to prioritise, is only identified, reviewed and challenged some way into the project. Another implicit assumption emerged in relation to using Acorn household Data as the primary source for market insights about customers. However, during workshop discussions we identified that the primary customers for some services are likely to be intermediary agents rather than households, e.g. architects and builders for planning and building related services. Additional market and internal customer service Data was also relevant and important.

By contrast, at InfraDig no inquiry logic was noted. Instead, for the engineers, questions of how to improve maintenance seemed settled. Instead, they were pursuing what can be described as a benefit realisation logic – to improve maintenance planning, scheduling and fieldwork by providing Data to maintenance practitioners. This benefit realisation logic too was left quite broad and vague. For example, it was not clarified how the Data being collected would

be used to generate maintenance improvements, by whom in which contexts, using which tools, and by solving which particular problems. In particular, it was not clarified how using the Data provided would improve particular practices and outcomes to crystallise particular benefits.

On reflection, pursuing and making such maintenance improvements seems likely to involve at least some inquiry. Practitioners will need to establish which areas of maintenance to prioritise for improvement using Data. However, this illustrates how inquiry might be embedded in action for practitioners and left implicit. It also highlights the potential interdependence of practitioner inquiry and realising that related practice improvement benefits from related insights. The observation that a separate, though inter-dependent, benefit realisation logic is also being pursued represents the first refinement to the explanatory framework, as illustrated below. It is reflected as Improved Practice and highlighted with a coloured background.

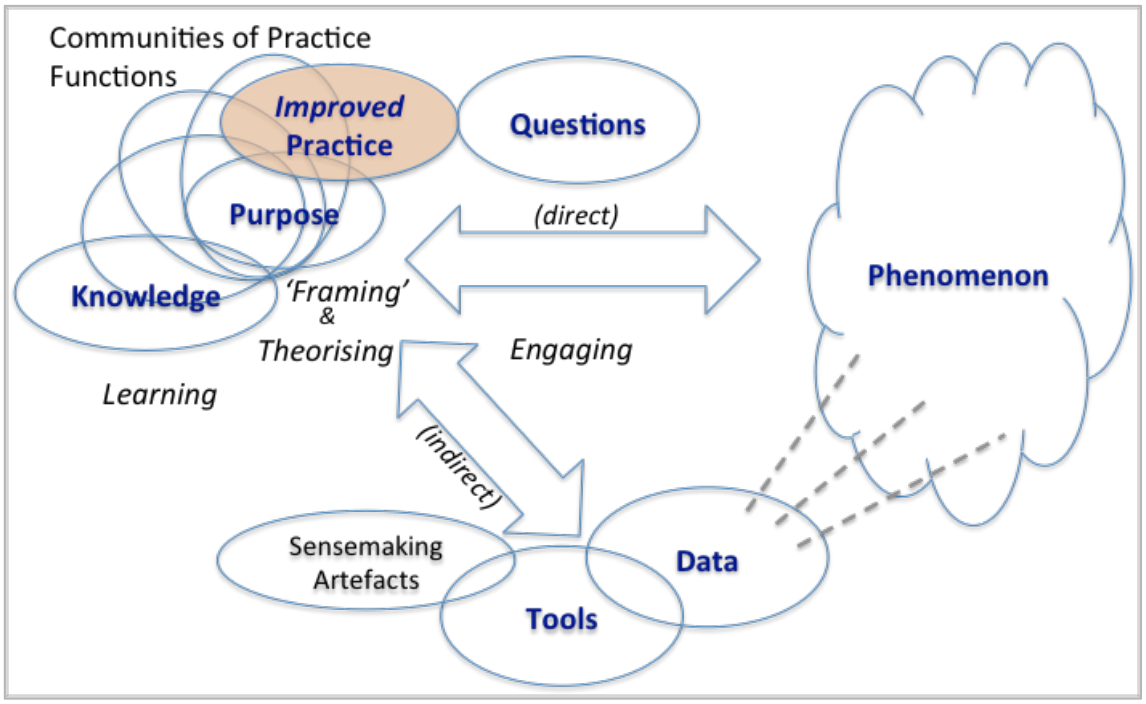


Figure 7-2 Benefit realisation framework refinement

Questions and inquiry

Research Question clarity and refinement has been a recurring focus during my research, and to frame and present my observations. It has been a source of reflection and supervisory discussion. It was also identified early on as a likely element of any inquiry framework, so something to pay attention to in the field. Turning to the two cases studied, they contrasted markedly in this respect.

At GoCouncil the sheer *abundance* of Questions generated during market insight workshops was striking, posing a prioritisation challenge. Questions were used to draw the attention of other participants to particular features about customers or the business model, prompting discussion and clarification. The abundance and variety of Questions sometimes reflected a lack of clarity about the Phenomena and related features of interest. However, this also reflected a lack of clarity and definition of the particular business proposition under discussion. Seemingly obvious Questions were also missed, e.g. in relation to pricing preferences, highlighting a lack of relevant commercial Knowledge and experience with business models.

By contrast, InfraDig was striking for the relative *absence* of Questions about maintenance, in spite of the existence of typical, illustrative operator Questions included in the COBie guidance (BIM Industry Working Group: 2011). Instead, for the engineers these Questions (and related Data requirements) were considered obvious and settled. However, it wasn't clear to me which Questions were being prioritised and addressed by the Data being collected. Nor was it clear to the operations handover planner, who highlighted that the Data would primarily be used to address audit and quality Questions. Within the asset Data team, the Questions that were noted related to Data and Tool *implementation*, e.g. which datafields to populate, and project logistics, rather than to maintenance. See the participant interview summaries in Appendix D.4. The role of Questions is revisited when I consider Engagement in section 7.5.

7.4 Practitioners shape Data initiatives

This section shifts focus from the Purposes and logics being pursued to address WHO is involved in pursuing and enacting them, i.e. participants. As already highlighted in the introduction and literature review, the typical logics underpinning such projects leave the WHO implicit. Also, as illustrated above, such logics often frame the project objectives and activities for participants. They inherently reflect the challenges and concerns of the particular Practitioner Groups involved, i.e. the WHAT and WHY of such initiatives. This section considers participants at the group rather than individual level, i.e. the Practitioner Groups involved. They are considered in terms of their role in realising the overall project inquiry and benefit realisation logics, their influence or control over critical resources, e.g. Data, Tools and staff, and in terms of their relationships and interaction with other Practitioner Groups involved.

Given its focus on Practitioner Groups, the CoP framework (Wenger: 1998) was important in informing these observations and findings. The comparative analysis of in-depth participant interviews is also relevant (see Appendix D.4). Given its scale, the number of different Practitioner Groups involved and the mapping Data collected, the InfraDig case looms large in this section.

7.4.1 Data initiatives as boundary spanning efforts

As already highlighted in the InfraDig case description in Chapter 6, several different Practitioner Groups are central to the InfraDig data initiative: Design Engineers, the Asset Data Team and Document Controllers. In addition, construction Contractors more broadly are closely involved, while the operators represent the ultimate Data users (with handover to them handled via the operations delivery team). The IT function also supports the hosting and availability of the core eB solution enabling Data collection, though the Asset Data Team superuser is more directly involved. These groups are shown below, with the groups central to the InfraDig Data initiative highlighted in blue, while other peripheral groups, i.e. less directly involved, are indicated with dotted bordering and in beige.

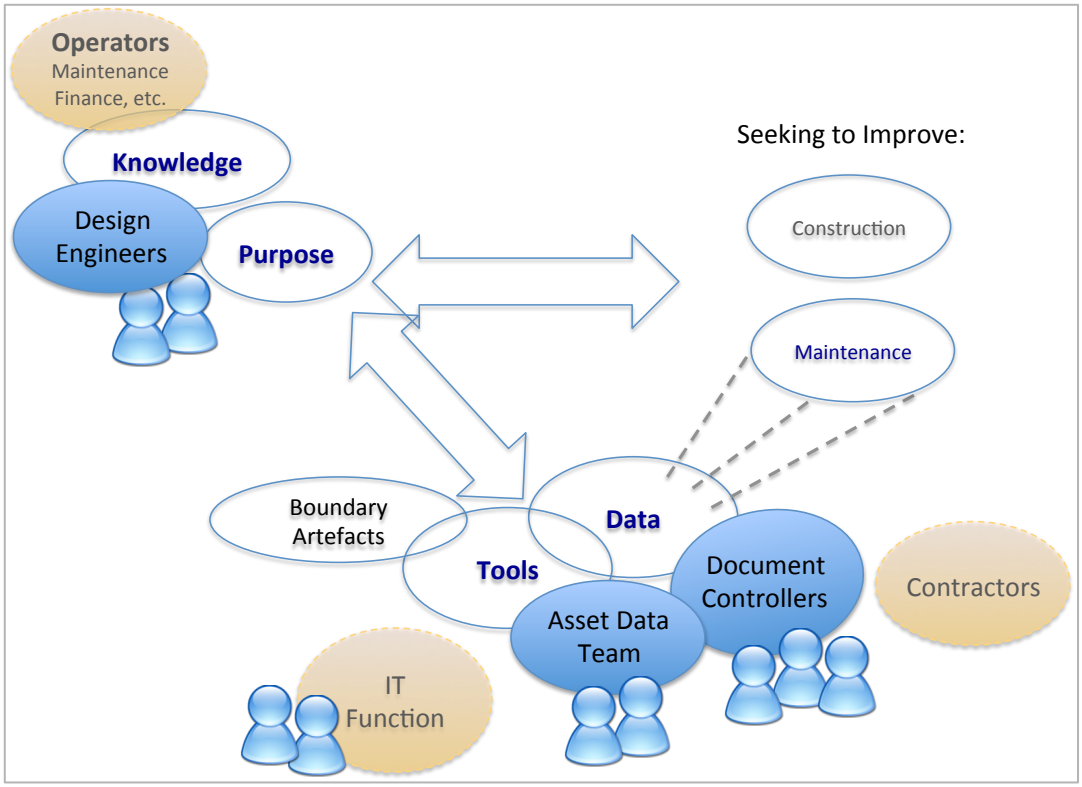


Figure 7-3 InfraDig Practitioner groups in relation to the data initiative

Figure 7-3 illustrates how some groups are particularly closely associated with specific framework elements, e.g. the asset data and document controller teams with Data and Tools, specialising in particular roles and activities:

- Data design and related tool selection (Design Engineers and their Operator equivalents)
- Data collection (Document Controllers and Contractor staff)
- Data management and related quality control (Asset Data Team)
- (Planned or intended) Data use (Operators' maintenance teams)

Such specialisation creates a need for coordination and alignment. Indeed, at InfraDig I noted considerable boundary spanning effort to align these Practitioner Groups, relying on the efforts of particular individuals or Boundary Spanners, although these participants framed this activity as communication and relationship management. Some boundaries noted occurred between different Practitioner Groups within InfraDig, often situated in different teams or functions aligned to particular practices (e.g. engineering, finance, IT). Other

boundaries noted were more organisational in nature, even when the Practitioner Groups could be identified as belonging to the same broad community of practice, e.g. design and construction engineers all belonging to the same engineering practitioner discipline but based in different units or organisations. The InfraDig case description also highlights imperfectly aligned interests in their contractual arrangements.

While the GoCouncil case involves fewer Practitioner Groups and clearly defined or specialised roles, they are nevertheless apparent, as well as the related Boundary Spanning effort. For instance, Marketing takes the lead on providing Acorn Data and related Data analysis to other functions (although some functions do collect their own Data as well). Acorn Data Knowledge and its management also reside with Tanya, based in the corporate team, with a social care practitioner background. There is a clear consciousness of working with various functional teams, and the broader business model initiative is recognised as organisation-wide. This is reflected in Albert's appointment from Planning, and the Deputy CEO's involvement in driving the Engine Room initiative.

On reflection, even my own research represents a Boundary Spanning effort. This first surfaces during my literature survey when reviewing knowledge management and situated learning literatures, coming from a grounding in and relative familiarity with the IS literature and discipline. It surfaces again during research design formulation, in terms of access to the social Phenomenon of interest, and identifying ethnographic immersion as the best approach to do so. This immersive researcher participation represents a form of peripheral participation (Wenger: 1998). Certainly I was conscious of being peripheral and new to the particular contexts of marketing and engineering. An example was learning about the particular Tools or systems and Data relevant to managing and maintaining engineering assets, and Acorn household Data, though I was familiar with Data, Tool and systems concepts in general.

7.4.2 The importance of Groups' domain Knowledge

Relating these Practitioner Groups to the wider framework and logic of inquiry, Data use and benefits realisation, highlights that these Practitioner Groups contribute different domains of Knowledge to the effort (which the effort may enhance). They also start with different levels of domain Knowledge in relation to the Phenomenon and other elements and activities (e.g. Tools).

In the case of InfraDig, design engineers bring considerable domain Knowledge relating to *constructing* infrastructure assets. Indeed, further specialisation was noted within engineering for particular asset types, which were classified into functional groups, e.g. power or civil assets or classes. The asset data team, by contrast, contributed considerable Knowledge on how best to capture, track and manage related asset design documentation.

In the case of GoCouncil, Marketing is clearly contributing domain Knowledge about obtaining and analysing market and customer Data, i.e. market and customer research Knowledge. Functional units contribute Knowledge about the particular services being designed or revamped, although, as highlighted in the case description, their level of commercial Knowledge about these was considered inadequate. Functional team Knowledge was also considered inadequate in relation to financial aspects, (although this could be argued to be a different domain of Knowledge, which the finance team might contribute). Indeed, the lack of necessary commercial Knowledge was recognised by GoCouncil and they had planned to provide training on how to build and use business models to address this.

Boundary Spanners (Wenger: 1998) in both cases had some experience of other Practitioner Groups' domain Knowledge. They drew on this Knowledge to facilitate, coordinate and clarify between Groups, in addition to trying to align them. In the InfraDig they were mainly from engineering, as a primary reference discipline, while they had some experience of asset maintenance (although not always in an infrastructure context). One Boundary Spanner had a good understanding of both engineering, and managing asset Data and related Tools.

By contrast, at GoCouncil Boundary Spanners instead had a generalist outlook and a reasonably wide experience within the council context, rather than deep specialisation in particular domains, activities or roles.

Relating domain Knowledge to inquiry elements highlights gaps

Mapping these different Knowledge domains to framework inquiry elements reveals that the central Knowledge domain relates to the Phenomenon of interest. It is also the primary focus for generating new Knowledge within the Data initiative.

In the case of InfraDig, the Purpose and envisaged benefits point to the importance of asset maintenance domain Knowledge. However, the dominant Knowledge domain within the data initiative is Knowledge about asset *design* and *construction* rather than about asset *maintenance*. This is embodied in the design engineering Practitioner Group driving the initiative. The Data collection effort seeks to capture this Knowledge about design and also Knowledge about the *construction* of these assets from Contractors building them. This is represented by the design and as-built documentation collected. Considerable domain Knowledge is also noted in document controller and asset data teams in relation to Data – in terms of how to collect, organise and manage those Data.

The Asset Data Team brings considerable Knowledge to bear in relation to the Tools (eB and Excel) that are used to collect and manage the asset Data. In the background, the IT team also contributes Knowledge about Data and Tools, mainly in relation to integrating and leveraging wider corporate Data and Tools, as well as hosting them in a controlled IT environment. However, no Data and related Knowledge is systematically collected or captured about existing maintenance practices within operators. To the extent that cost-effectiveness and greener maintenance are also encompassed as Purposes (with related Phenomena), finance and sustainability Knowledge domains are also relevant. Yet these are not reflected in Data collection or in the ultimate user Practitioner Groups involved.

In the case of GoCouncil, customer and market domain Knowledge are central to inquiry and benefit realisation. To the extent that this is a Data-led project, such Knowledge is chiefly encoded in the Acorn household Data that is being used to shed light on likely customers and their preferences (the Phenomenon of interest). Tanya's Acorn Knowledge emerges as critical. Marketing's broader domain Knowledge about additional customer Data collection complements this. Their prior Knowledge of residents and potential customers, gained through previous Marketing interactions, also emerges as relevant, though remains tacit. GoCouncil recognises its lack of commercial Knowledge, even in the finance function, and the need to address this in order to achieve their overall objective, yet the planned related training programme is dropped under cost and time pressure.

Across both cases, the domain Knowledge areas identified either relate to the primary Phenomenon of interest, or to the enabling activities, techniques and Tools used. In addition to highlighting potential gaps, such analysis of Groups and their domain Knowledge highlights how this is both embodied in specialised Practitioner Groups, and explicitly captured or reified in Data artefacts. However, in neither case are specific domain Knowledge or learning goals made clear; nor is domain Knowledge or emerging learning explicitly monitored or managed.

My research highlights how a formal research design makes relevant domain Knowledge and method related elements explicit. These are kept under formal review with a supervisory panel. Indeed, my panel was selected in part for their domain or method Knowledge. Planned domain and method contributions and intended dissemination are also made explicit and discussed during later reviews, e.g. conferences to present and refine ideas. This can also be argued to represent an academic equivalent of benefit realisation (i.e. impact in UK academic jargon).

7.4.3 Separation of Data collection and use

Considering Practitioner Groups' roles in relation to particular framework elements also highlights the separation of data *collection* activities from Data

use and *Engagement* activities (and both activities from Data management and organisation). This seems to reflect the specialisation already noted, as well as pragmatically addressing logistical and scale requirements.

This is most clearly illustrated in the InfraDig case, given the scale of its Data collection effort and physical location of Data collectors. The centrally based design engineers are primarily responsible for capturing Data into design documents. Contractors are responsible for updating these or 'red-lining' them to reflect 'as-built' differences. Document Controllers locally coordinate this activity and are based at construction sites, capturing the updated documentation into the asset data Tool. The Asset Data Team centrally manages the documentation Data captured into the Tool (both by design engineers and Contractors). The team also manages the Tool and performs basic Data quality checks.

Physical separation of these activities reflects where the construction activity takes place and the activity to update the related documentation. Separation from ultimate use is more complex. Firstly, it is separated in time, as maintenance of the asset and the use of the Data collected for this purpose will only commence after handover. Secondly, maintenance planners are also physically separated in that they reside in the ultimate operators, although the actual maintenance activity is likely to be contracted out, possibly even to the same Contractors that were involved in the construction. However, these contracts are not yet awarded and the involvement of the same Contractors is by no means certain, nor is it clear how much of a role Contractors will have in planning the related activity. The original design engineers are unlikely to be involved or available during the maintenance phase. Similarly, the Asset Data team will be disbanded at the end of the project and Operators will take on Data management. Such physical and temporal separation is reinforced by the addition or introduction of a contractual relationship between some Practitioner Groups. The above points to potential conflicts of interest and a lack of contractual alignment between Contractors, Infradig and the ultimate Operators, which is dealt with in the next section.

The GoCouncil case also reveals a clear separation of Data collection activity and Data use in relation to Acorn Data. Acorn Data is compiled and made available by a commercial organisation, using UK census data, collected to inform public policy and related decision-making. By contrast, in Helga's focus group effort there is very little separation of those involved in Data gathering and use. *Similarly, I was involved in all aspects of Data collection, organisation and subsequent interpretation.*

The Practitioner Group specialisation and separation of Data collection and use represent important refinements to the provisional explanatory framework. They also underline the importance of Boundary Spanning activity, which represents a further refinement.

Figure 7-4 shows additional Practitioner Groups specialising in Data collection and related Tools, that are separate from those Groups related to Data use, practice improvement and benefit realisation. Boundary Spanning is also introduced, especially in relation to spanning this separation. Refinements are shown in colour again to highlight them against the provisional framework.

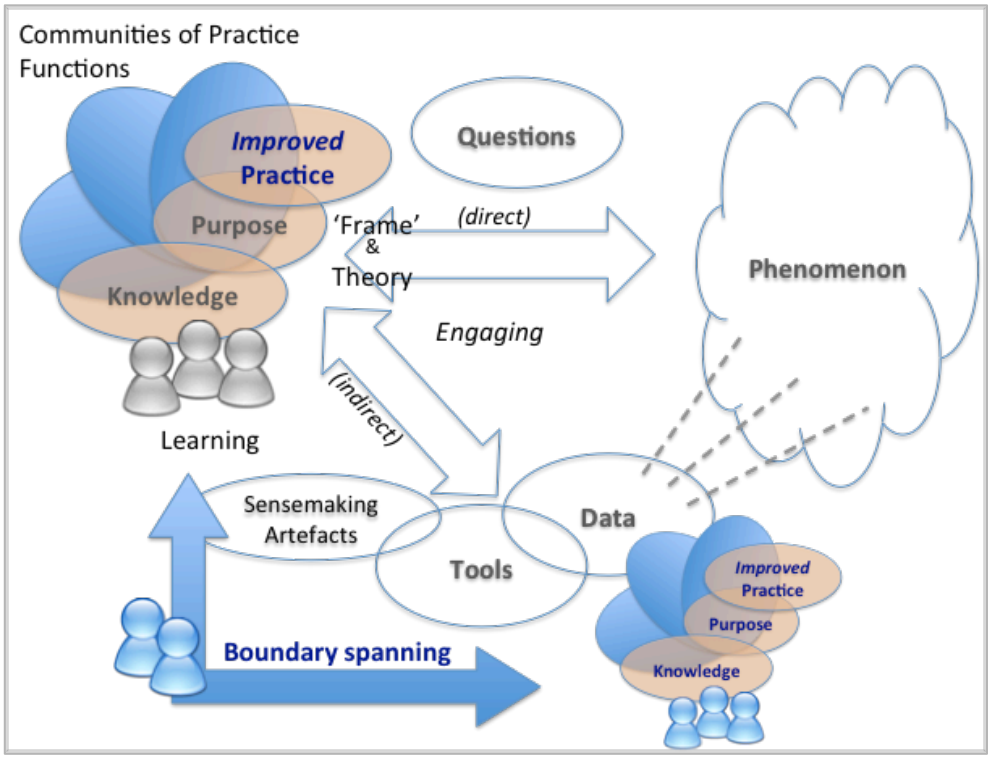


Figure 7-4 Specialisation and boundary spanning framework refinements

These separations and specialisation can have important implications for interpreting and using Data, and are revisited in section 7.7.5 below.

7.4.4 Challenges experienced aligning various groups' activities

Before turning to the question of Data Use and sensemaking, this section touches on the challenges observed in aligning the various Practitioner Groups involved. Rather than focusing on individual leadership, which is also likely to be important, this section examines Practitioner Groups' relative ability to influence or control other Practitioner Groups involved. This influence is assessed in terms of the initiative's Purpose, resources, economies of meaning, and activities or practices pursued. Various strategies employed by 'leading' Practitioner Groups to achieve such alignment and realise benefits are also observed.

InfraDig – Influencing Data Collection

At InfraDig, the overall asset data initiative is clearly led by design engineering and becomes identified with this Practitioner Group. The positioning of the Asset Data team, tasked with organising and managing the Data collected, within engineering reflects this.

However, Document Controllers are based on site at the construction project and report on a day-to-day basis to the on-site project field engineer, who reports up the project function via the local project manager, i.e. not to design engineering. While these are employed by and represent InfraDig on site, they work with equivalent Contractor staff, responsible for obtaining and providing the necessary documentation from their engineers so this can be captured or loaded into the eB system. As already highlighted above, and in the InfraDig thick description, Contractor and InfraDig interests are not necessarily aligned. This is imperfectly addressed in current contractual arrangements, which still reflect a relative emphasis on delivering the physical infrastructure, rather than related Data. While Engineering can exert line management control over their design engineers and asset data managers, they have to internally influence their project field engineers and Document Controllers, hoping these in turn will

influence their Contractor equivalents. Nevertheless, the daily reality observed, was that InfraDig and Contractor Document Controllers worked well together locally to complete Data collection.

Alignment is partly addressed through Boundary Spanning activity, especially by the chief engineer, who tries to attend the periodic project field engineer forums, using his seniority and management position, as well as CEO support for BIM, to secure access. Some quality assurance activity, on documentation submitted, is also performed by a small quality control team, which is situated in the design engineering function. This team carries out spot checks and reports on Contractors' compliance with contract requirements to provide as-built documentation. This is reported to local InfraDig project managers and project field engineers to address with their local Contractors.

Local project managers and their field engineers have considerable discretion to sign off Contractor completion as adequate at a local level, and this is often where pragmatic compromises on Data quality crystallise, perhaps in favour of physical construction priorities. This plays out during formal contract review meetings and the deliverable milestone sign-off process. For some of these the operations handover team are also involved and represented. To the extent this team doesn't include asset data representatives though, Data issues may not be raised or prioritised. To address this risk, the chief engineer is constantly making the case at field engineer forums for the equivalent importance of completing the 'virtual' infrastrucure and the physical infrastructure. However, this represents a significant shift in field engineer and project manager framing of their responsibility and practice.

As an additional strategy to mitigate this risk and improve alignment, the quality control team within engineering also establishes a Document Controller forum, recognising the potential to create a community of practice. However, this Group seems to frame their role as administrative and appears to have very limited understanding of the content of the Data they are collecting, as well as the significance and importance of those Data for maintenance. The intended KPI reporting on asset data collection offers a further mechanism for monitoring,

influencing and aligning such activity. The chief engineer will formally review this reporting with the InfraDig management team at management meetings. The InfraDig CEO also wants InfraDig to be seen as leading practice in relation to the government's BIM objectives. This is reflected in support during these meetings, as well as agreeing to engineering resources to help realise these objectives, i.e. funding for the asset data team, quality assurance activity, supporting systems, etc. While the chief engineer would also prefer stronger contractual provisions around Data document quality, these contracts were formulated and agreed at the outset of the project, before the data initiative had gained traction.

InfraDig – Influencing maintenance practice improvement

The Asset Data Team's ability to influence and align Operators proves even more tenuous. InfraDig's relationships with Operators are also essentially contractual, and considerable effort is directed to agreeing related *formal* asset Data requirements. However, here InfraDig is the supplier and the Operators are the client, so InfraDig's influence is limited. While Boundary Spanning effort is expended here, mainly through asset data forums (at a relatively senior level) with Operators and their engineering functions, these seem to have limited success. They mainly appear to focus on trying to achieve consensus between Operators as to Data definitions and classification approaches.

One Operator was still being set up and had yet to appoint an equivalent Asset Data Team for InfraDig to engage with. They also had less invested in existing systems and approaches, i.e. institutionalised, so were proving less resistant to adopting InfraDig suggestions. No contact between the wider asset data Practitioner Groups at InfraDig and Operators was noted. In contrast to Document Controllers, there was no specific initiative to foster a similar CoP and promote shared Data related practice improvement.

GoCouncil – influencing the use of market insights for business models

GoCouncil had fewer groups and organisations involved, no contractual arrangements and little physical distance to contend with. Nevertheless, Marketing also experienced alignment challenges. Marketing has no direct

influence over other functions using them, or their market and customer research; they are essentially peer groups. Their involvement is also explicitly framed as support and coaching, although they do insist that explicit market or customer research needs to be coordinated through them. They argue this is to ensure coordination, consistency and achieving synergies. However, Marketing's representation on the board and at council meetings provides visibility and a platform, which it uses to motivate for the market insights initiative, to keep the board informed of progress and demonstrate the value being added. In due course, it also has a voice at this forum to challenge business proposals put forward that are insufficiently informed by customer and market research, and testing to ensure they are likely to be viable and successful.

In addition to Marketing's representation on the board, it is also responsible for internal communications, which represents a powerful promotional tool. Marketing also seemed adept at collaborating with the wider transformation team, involving Albert on the project team and securing the sponsorship of the deputy CEO for the Engine Room. However, while their coordinating role and 'ownership' of market research was not directly challenged, subtle challenges to their influence were noted. The most notable example was when some board members posed the question: how much market insight is enough?

7.5 Iterative, progressive clarity through Engagement

I now turn to the heart of the explanatory framework and the activities noted that seem most closely involved in pursuing insight. The section shifts our focus from WHAT is being pursued, to consider HOW insight is generated. This activity is tightly bound with WHO is involved in these activities.

Wenger (1998) posits the importance of three broad sets of activity to promote emergent learning: Engagement, Alignment and Imagination. As already described, InfraDig highlights the need to align different stakeholder groups around the overall purpose of a data initiative, and to ensure incentives are aligned. InfraDig also provides evidence of engineers *imagining*, through their 3D tool development and an initiative to re-imagine maintenance in 2025.

However, these activities and related challenges are common to many complex initiatives spanning different stakeholder and Practitioner Groups. It is the level of Data *Engagement* activity that seems distinctive about Data initiatives, and which is examined more closely.

7.5.1 Data Engagement and use as critical to new insight

Considering Data Engagement, the first related questions to arise are: WHO is doing the Engaging and with WHAT? While the explanatory framework anticipates direct Engagement with a Phenomenon as well as Data about it, both of the cases are ostensibly Data-led and use Data as their starting point. GoCouncil has readily available Acorn Data, while InfraDig is collecting, digitally capturing, organising and relating various elements of design and as-built documentation. However, while both cases are Data-led in their initiatives, they are characterised by very different kinds and levels of Engagement with Data.

Those who Engage learn

At GoCouncil, Tanya leads the Data Engagement effort. She summarises Acorn Data for various target household categories to make it easier for staff less familiar with Acorn Data to engage with. She geographically maps some of this information visually to further aid Engagement and sensemaking. She also makes the more detailed Data available, offering help and coaching. The workshop sessions introduce Acorn Data to other Practitioner Groups, to demystify it and encourage its use. In the process of summarising the Data, she identifies patterns and generates insights, e.g. gardening interests, which she then raises with the Leisure team as a potential service opportunity. While Engagement with the Acorn Data was also clearly evident within workshop exercises, there was less visibility of subsequent Engagement by functional teams. This was not monitored by the project and I had less access to these teams.

At GoCouncil there was a readiness to Engage *directly* with potential customers (the broad Phenomenon) to complement their Acorn Data Engagement. Given her long Marketing experience and familiarity with Acorn Data, Helga quickly

recognised Acorn's limitations to address intermediaries; OtherCouncil also recognised its measurement limitations to address deprivation at the individual rather than household level. In their case, they used relevant Data available from their internal systems to supplement Acorn Data. These decisions seem pragmatic, e.g. OtherCouncil's ready availability of such complementary Data and skills to extract it. At GoCouncil, relevant intermediary Data were not readily available, while it was relatively easy for Helga to arrange a focus group with some intermediaries.

By contrast, at InfraDig lots of Engagement was noted with the *collection* and *organising* of Data, rather than Engaging in *using* it to explore the maintenance Phenomenon and generate maintenance insights. This reflects the various, specialised Practitioner Groups involved, as per Figure 7-3. It also highlights the *absence* of maintenance Practitioner Groups, who might Engage meaningfully with these Data to improve their practice. It is therefore not surprising that no new maintenance practice insights were noted during fieldwork. Such potential Data Engagement is effectively deferred until after handover.

For InfraDig, even deciding how best to organise and present the Data to Operators emerges as challenging and contested. On the one hand, these tensions reflect the different operator views about which Data elements are likely to be relevant. The Operators' concerns about integrating the Data into their existing operational systems reflects their focus on current operational use rather than exploratory analysis, i.e. data exploitation rather than exploration. The tensions noted about how best to organise and present the Data reflect how Data organisation is optimised and encoded within tools and related data structures with particular contexts and use in mind. This is considered in more detail in section 7.6.

On the other hand, InfraDig is proposing a Data artefact based on what they think will be most useful for maintenance. InfraDig relies mainly on the prior maintenance experience of Boundary Spanners such as the chief engineer and some of the consultants involved. Some prototype tool design and related *imagining* (Wenger: 1998) is also noted, e.g. the 3D tool design and

demonstrations. However, none of these addresses the particular operator maintenance contexts in view using Data specific to these contexts.

Data Engagement not a particular focus

Rather than facilitating Operator Data Engagement, InfraDig participants frame the challenge as a problem related to agreeing Data *requirements* with Operators. InfraDig don't recognise the inherently abstract and somewhat unknowable nature of specifying *future* requirements. For example, no efforts were noted to co-develop asset Data working templates or prototypes with Operators. Instead, relatively abstract requirement discussions are held, often between fairly senior engineers, asset data specialists and consultants, rather than staff engaging on a day-to-day basis with asset Data for maintenance. The seniority of attendees also reflects the mainly contractual, rather than exploratory and collaborative, nature of these discussions.

A natural opportunity for early Operator Data Engagement is also missed during infrastructure handover (still underway). Sections of the infrastructure, e.g. some sites, will be delivered in stages with some ready earlier than others. Related asset Data, also provided early, could facilitate testing and refining related requirements, e.g. in terms of content validity and completeness, presentation and integration. This could inform improvements to the remaining Data collection and organisation effort, but was not included in handover plans. This increases the risk that the Data provided may not be considered fit for purpose by Operator maintenance teams in due course. It may reduce confidence in the Data, familiarity with it and consequently reduce its use and related Engagement to produce maintenance insight.

On the one hand InfraDig misses several opportunities it has to refine what Data are collected, as well as the methodology, measurement and presentation approaches used. On the other hand, GoCouncil has little or no scope to influence the Acorn Data supplier in terms of what is collected, how this is done or its presentation. This provides a rich illustration of how different CoPs have different levels of influence over the data they use and how those data are organised.

7.5.2 Institutionalised practices frame inquiry and Data Engagement

InfraDig's focus on Data requirements rather than Engagement highlights an important loss of focus and emphasis on the inquiry Phenomenon and on realising the ultimate maintenance benefits. Instead, their focus is foreshortened to Data collection and producing a Data artefact as an end in itself.

Considering institutionalised practices at work offers one route to explain this. Institutionalised practices are an inherent feature of CoPs (Wenger: 1998) in order to get things done collectively and effectively. They often involve *economies of meaning* (see section 2.6) and associated reifications or codifications, for example in standard documentation, procedures and systems. While some of these reifications address local activities within a Group, others play an important role in coordinating effort across Practitioner Groups. These practices and artefacts frame activities and related sensemaking for participants.

Institutionalised practices and related Boundary Artefacts are particularly prominent at InfraDig, given the number of specialist Practitioner Groups involved. They were first noted in the asset reporting IT project, in relation to configuring the eB system, which represents a reification of both asset related Data, related Data organisation and Data management practices. Considering the IT project activity, it reflects the institutionalised practice of following a defined and well-understood project development process, with associated requirements' definition documentation. This documentation represents a Boundary Artefact with which to agree and coordinate development activity between IT development staff on the one hand and users on the other.

A similar 'project paradigm' was also a feature of the wider asset Data initiative (and indeed the overall InfraDig construction programme). The project contracts, design, as-built, and operational documentation loom large as important Boundary Artefacts, to coordinate activity between the project team designers and construction Contractors on the one hand, and with the infrastructure operators on the other (together with the handover plan). These have evolved and been institutionalised across the industry over many years

and many infrastructure projects, mainly between engineers, who bring considerable tacit engineering usages and experience to bear in interpreting them during use. These Boundary Artefacts and this 'project paradigm' provided a pervasive 'frame' for activity, which emphasised a bounded, short-term timeline and end – the production of project deliverables, e.g. asset Data.

Turning to GoCouncil, the dominant area of practice drawn on in their data initiative is that of market and customer research, situated within the wider Marketing discipline. Marketing's focus and activities reflected a Data orientation and they seemed to value both qualitative and quantitative approaches to collect additional relevant customer and market Data.

In my admittedly limited experience, research Data Engagement seems similarly enabled and constrained by institutionalised frames, practices and Tools within particular discipline areas. This was most visible in terms of norms of what is considered good research practice, which techniques and Tools to use, and which presentational formats are acceptable.

7.5.3 Questions can focus Data Engagement on the Phenomenon

Turning to the role of Questions during Engagement and inquiry, the cases reveal how these can guide both direct Engagement with the Phenomenon or Engagement with Data about the Phenomenon of interest. At InfraDig they even arise in relation to the logistics and tools of inquiry. Indeed, Questions are revealed as a flexible social focal and engagement device, particularly during productive discussion, drawing attention to particular features of interest, anomalies, or patterns.

However, Questions seem most helpful for inquiry when they help clarify Data requirements and relate them to a particular Phenomenon. *While the GoCouncil 2x2 mapping exercise first brought this into focus, identifying how Questions related to available Data, it also emerged reflexively. Research Questions were key to demonstrate alignment and consistency of research design elements. However, in contrast to my research, neither case maintained an ongoing focus on the Question(s) they were addressing, in order to clarify how the Data*

collected would be used or analysed to address them. While GoCouncil did include explicit Questions in their original board market insight proposal, these Questions and the related analysis weren't actually used to formally manage or prioritise customer and market Data analysis.

At InfraDig, COBie Questions were not widely communicated, well understood or prioritised, nor used or revisited periodically to facilitate clarification and refinement of Data collection. How asset Data collected relates to maintenance and which Questions are in view, is left implicit and untested, deferred until after handover. Indeed, when challenged about the relative absence of Questions in their discussion and focus, it emerged that they think many of the COBie Questions are settled for them, possibly based on their prior experience; rather it seems more about convincing Operators. Given the range of potential stakeholders identified at InfraDig, all with an interest in using the Data and in achieving cost-effective maintenance, one can foresee a range of Questions arising for different Practitioner Groups. InfraDig engineers are unlikely to be able to anticipate all of these to ensure the Data collected can address them all effectively.

7.5.4 Iterations of clarification

Here my research experience contrasts considerably with both cases studied, where I observed relatively little progression and clarification, or iterations of inquiry and engagement with Data. The process of documenting a clear inquiry logic in the form of a research design, and its ongoing review, highlighted various implicit assumptions and areas requiring further clarification or specificity, e.g. the Phenomenon, unit of analysis, etc. Progressive clarification and refinement of the inquiry logic and Phenomenon has been a feature of my research, with greater clarity emerging from each iteration of Engagement with the literature, Data collected and during writing, reflecting, discussing and presenting my observations. This process was characterised by increasing confidence in the research design and its stability, and confidence in generating useful insights. Indeed, there has also been increasing confidence in the observations and insights that actually emerge themselves.

The relative lack of progressive clarification noted in cases may in part be due to their inquiry logics remaining largely implicit and unexamined. It may also reflect the relatively short period of fieldwork immersion. By contrast, my research journey and progression has occurred over several years. Evidence for the possible need for considerable time and iterative refinement, was also found in my interview with OtherCouncil. Even though they were relatively clear about their Purpose and focus on the deprivation Phenomenon from the outset, it nevertheless took them approximately two years of refinement and experimentation to achieve a flexible and stable deprivation index and related Dataset that they felt was useful across different Practitioner Group contexts.

Expectation of Question progression

At the research design stage, ahead of fieldwork, I had an expectation of seeing Question progression as an indicator of learning and clarification. This was based on Blaikie's (2010) characterisation of research Questions moving from What, through Why to How questions with greater understanding of a Phenomenon, as well as Tsoukas' (2005) argument for increasing knowing being the ability to make increasingly fine distinctions about a Phenomenon. Certainly I have noted that my research Questions have shifted and become more specific, as the research has progressed. This is evidenced in various documents produced at different stages, as shown in Table 7-1.

Table 7-1 My Research Question progression

Stage	Research Question
Research Proposal stage, based on limited reading – May 2010	**Are the theory constructs and hypotheses put forward by Marchand et al. (2001) supported by the evidence of experience in the context of CRM, BI and KM systems? Which additional constructs and hypotheses are important to include in a more comprehensive theory that seeks to address some of the shortcomings highlighted in the problem description section?**

1st Review stage, based on initial literature review – October 2012	**How do [customer*] analytics project teams use questions to draw insights from [customer*] data?** * At this stage I was considering limiting the research scope to a customer analytics context to make the scope more manageable.
2nd Review stage, preliminary data analysis – January 2014	**How do teams 'frame' data initiatives to generate insight from data?**

Stage	Research Question
3rd Review stage, initial drafting of thesis findings-conference papers – June 2015	**What is revealed, by using a CoP lens for Data Initiatives, as important for enhancing insight and related benefits?**

The progression shown reflects a clearer view of my Phenomenon of interest, the particular dimensions of interest, as well as an improved grasp of existing theory, concepts and language used within the academic IS literature. A certain path dependency is also discernable – progressing from one question to the next, having addressed the earlier one. Questions tended to occur in clusters (sometimes in subordinate relationship to a broader Question). Some of these reflect the broad research process too, e.g. What does the literature have to say about the Phenomenon, or specific aspects of the Phenomenon? How best to research the Phenomenon? What Data to collect? How best to analyse the Data? Considerable prioritisation effort was also required, particularly at the framing and design stage, in terms of which Questions would be most interesting (and feasible) to pursue. Even the write-up stage involved selecting which research Questions to focus on to best frame the contributions emerging from my observations.

Turning to the cases studied, a lack of making Questions explicit precluded tracking such changes and related learning or clarification. Unsurprisingly, no Question progression was noted in either case, although, as highlighted earlier, this may require a much longer longitudinal study.

An absence of clarifying, testing and refining the Questions in relation to the Phenomena of interest and Data, also increases the risk that critical Data may not be collected to address some Questions (e.g. in relation to energy usage at InfraDig), which may require further (expensive) Data collection at a later stage. It also precludes prioritising Data elements for collection and Engagement. For example, at InfraDig this could inform Data quality efforts, and facilitate easy extraction, reporting or presentation design. At GoCouncil, the 2x2 Question-

Data analysis was indeed intended to help prioritise further Data collection by a resource constrained Marketing team.

The iterative nature of Engagement represents a further refinement to the explanatory framework and is illustrated in Figure 7-5 by introducing various circular arrows.

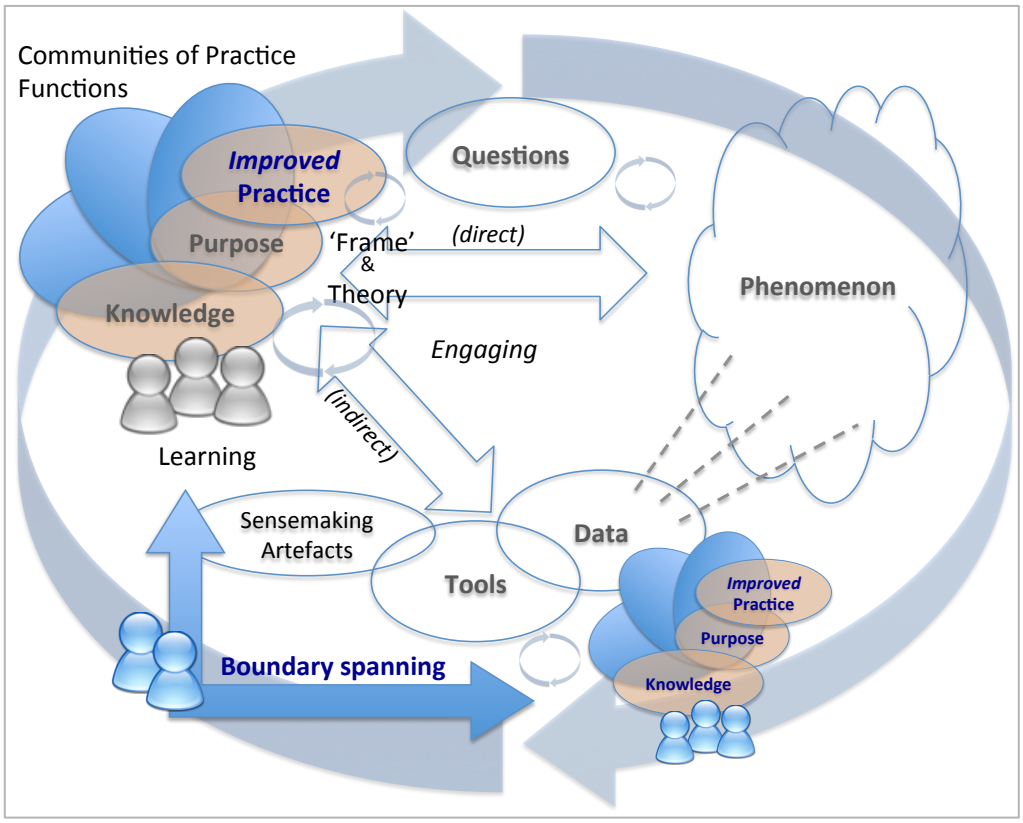


Figure 7-5 Iterations of Engagement as a framework refinement

Two broad arrows represent broad iterations of Engagement at the overall project level, with smaller versions indicated in relation to particular elements. This seeks to reflect that learning can occur at various levels. For example it may include 'reframing' the Phenomenon, in addition to refinement of the existing Knowledge about the Phenomenon of interest. It can also trigger new ideas as to Purposes or opportunities for Practitioner Groups, new Questions about the Phenomenon, identifying new Data dimensions and (Data) Engagement ideas. The complex, iterative 'dance' depicted highlights how Engagement can lead to second order or double loop learning (Argyris & Schön: 1978) in addition to the original learning objectives in view.

7.6 Data Artefacts and related Tools used

This section focuses on the Data and related Tools used. These are located within particular Practitioner Group contexts. So, as a starting point, the section summarises, and compares and contrasts the cases in terms of their use of them, also considering my Tool and Data use reflexively. The section then goes on to highlight several related observations.

7.6.1 Overview of Data Artefacts and related Tools used

At GoCouncil, Acorn Data looms large. These Data are extensive and detailed, and generally made available in an Excel workbook format. However, Excel isn't used for extensive Data analysis or to aid inquiry by applying analytical techniques or Data processing. Instead, most of the practitioners Engage with the target group summaries prepared by Tanya, essentially a related Sensemaking or Data Artefact. These summaries reflect a prioritisation of household groups and what Data about them is likely to be most relevant to participants (i.e. filtered). They include some Data visualisation, highlighting where these households are prominent within the council jurisdiction on a map. Tanya uses a Geographical Information System (GIS) system to achieve this.

GoCouncil's use of Data and Tools contrasts with OtherCouncil. In addition to using Acorn Data, they use considerable data from their internal operational systems. Their Data Engagement was pursued using fairly basic Excel and SQL type Tools. At GoCouncil, data fragmentation across a number of systems, varying data quality, data privacy, and a lack of ready SQL and data skills, are cited for not using internal systems data. Instead, Marketing motivate for a new CRM system.

At InfraDig an asset Data Artefact is being created for Operators to represent a 'Virtual' infrastructure, i.e. an abstraction of the physical infrastructure, to facilitate sensemaking about the physical infrastructure and related analysis. While essentially a collation of considerable design and as-built documentation, in various formats (Microsoft Word documents, CAD drawings, etc.), these nevertheless represent a form of codified Data, even if less structured than

more traditional Data within databases. InfraDig uses eB as a repository for collected Data, to coordinate submission by various users across multiple locations, and to manage mapping and interrelating Data. Data Engagement involved using Excel and SQL, reflecting inherent limitations of the proprietary eB query language and very limited visibility of the eB Data structure (considered proprietary).

By contrast, the Tools that will be used by Operators' maintenance planners and teams to Engage with the Asset Data, are likely to be those they already have in place. Even where similar Tools to InfraDig are being used, these are not configured in the same way, reflecting local practices and Data usages. While actual Data Engagement is deferred, InfraDig develop a 3D Tool as a prototype or pilot, which uses illustrative Asset Data to facilitate a virtual tour of the infrastructure, to spark ideas of how the Asset Data might be used. Essentially this seeks to promote *imagination* rather than engagement (Wenger: 1998), to inform creating a vision for maintenance in 2025, which will represent a further sensemaking artefact to engage with.

Reflecting on my Tool and Data use highlights various similarities and differences to the cases observed. My use of Excel, NVivo and Prezi, is similar to GoCouncil and OtherCouncil's use of end-user or individual user Tools. In the case of Prezi, I used it both as an exploratory Tool, plotting Data on a timeline, as well as for presentational purposes, creating various Sensemaking Artefacts to aid discussion with InfraDig during the group mapping exercise, and for colloquia. My use of NVivo to collate, organise, track and relate Data collected is more similar to InfraDig's use of eB to aid Data collection and organisation. While my use of NVivo, to classify and interrogate Data, is somewhat similar to InfraDig's use of SQL and Excel to check on Data Quality, it also went further. NVivo was used to aid reflection and actual Data Engagement, in listening to recorded data, annotating and actively analysing such classification, in order to generate insights from the Data.

Reflecting on the Data and Tool use noted and described, crystallises several important observations, which are outlined in turn in separate sections below:

- The socially constructed, evolving nature of Data in relation to Phenomena
- Recognising how Data can act as Boundary Artefact or infrastructure and the impact of their being used in this way
- How Tools simultaneously enable and mediate Engagement with Data, often based on reified, encoded knowledge

7.6.2 Data as a socially constructed snapshot of Phenomena

Building on section 7.4 and how practitioners shape data projects, InfraDig in particular brings into focus how Data are socially constructed during data design and collection, and how tools reflect institutionalised practice and related investment. First we consider how this is reflected in Data used, in terms of the views of the Phenomenon addressed and economies of meaning encoded, as well as how Data evolve and and are influenced by pragmatic considerations about what is available and easy to collect.

Data reflect different Purposes and particular dimensions of Phenomena

The Asset Data InfraDig collects and incorporates in its 'virtual' infrastructure Data Artefact, is necessarily a subset of what might be relevant about the asset or infrastructure for maintenance purposes. This reflects the knowledge of the engineering designers and institutionalised Operator thinking during requirements discussions. The InfraDig operations handover team also intend the asset Data to evidence adequate quality assurance checks of the infrastructure in readiness for handover, in case contractual problems arise subsequently. This represents an additional Purpose and is likely to require specific, additional Data or meta-Data elements to be collected for subsequent use and analysis. It also represents a shift or difference in focus, from the Data needed to address maintenance to those required for construction quality assurance. While both relate to the infrastructure being built as a broader Phenomenon, they represent different perspectives or views, focusing on

different dimensions or aspects of it. They can represent more specific, though related, Phenomena in their own right. In the light of COBie considerations, further perspectives or lenses can also be anticipated, for example in relation to cost and energy footprint.

In the case of Acorn Data, while they are initially collected for a broad set of public policy Purposes, rather than specifically to enable customer targeting, they are widely used for such targeting too, especially by Marketing teams outside the public sector in the UK. GoCouncil and OtherCouncil illustrate how these Data are essentially repurposed. Indeed OtherCouncil pragmatically supplement them with operational systems data in the case of OtherCouncil, while Helga arranges focus groups to capture intermediary Data.

Pragmatic incorporation of Economies of Meaning

InfraDig's proposed Data Artefact adopts and extends existing usages and standards of Data and document organisation, as well as related data collection practices. This continuity eases design and requirements discussion as well as Data collection between related organisations and Practitioner Groups, e.g. construction Contractors, InfraDig engineering and Data teams, as well as Operators. It is also likely to ease use and engagement by Operators' maintenance users, where it uses shared economies of meaning. However, different standards, usages, practices and systems within the various infrastructure Operators are also recognised, which will require either multiple presentation formats or a flexible structure and format, to aid its integration into operators' systems and encourage such local use. Indeed, several members of the Asset Data Team mention the need to standardise, or at least map usages, classifications and terminology between operators.

InfraDig's use of eB with its underlying, incorporated Data model, illustrates how Tools can provide a default or ready-made economy of meaning. This is richly illustrated during the KPI workshop discussion of fields available. The eB Data design reflects the original Purposes in view, design considerations and anticipated use when created, as well as subsequent refinement based on changes in response to changing customer requirements and feedback.

However, they pragmatically chose the system as the best available based on a selection exercise, reflecting broad areas of functionality thought relevant, rather than a detailed assessment of the underlying Data model.

Similarly for GoCouncil, Acorn Data encode or reify considerable institutionalised practice, mainly in relation to what Data are relevant to public policy decisions, as well as how to collect these Data and decisions as to the most appropriate or useful unit of analysis – households. Therefore, initial Data design considerations, selections and choices are not visible or transparent to InfraDig or GoCouncil, yet they implicitly frame or limit their view of Phenomena addressed, in terms of what is relevant to know and capture about them.

Tanya's creation of more accessible Acorn summaries for target categories, introduces further simplification and a filtered view for Functional users based on what she thinks is likely to be most relevant to inform business model design and evaluation. This reflects a trade-off between making the summaries and Data easy for users to engage with and potentially filtering out Data that may be relevant.

My own collection, organisation of Data and their use also reflect economies of meaning and institutionalised practice. My Data reflect both broad research methods' economies of meaning and practice, e.g. source coding of participant data and coding analysis, as well as more specific disciplinary considerations and ideas, e.g. using CoP ideas as a priori codes.

Data collection and selection is pragmatic

GoCouncil's use of Acorn Data reflects that it already holds a licence to use these Data, so they represent a free resource for the market insight initiative. It also doesn't have resources to undertake extensive Data collection about household preferences in relation to current and proposed services. Only limited focus groups and consultations might be possible. While InfraDig has greater resources available to collect Data, it nevertheless adopts or leverages existing processes, project resources and collecting standard design, as-built and Operating Manual (OEM) documentation, rather than embarking on completely

new data collection. Even modest approaches to use tagging of relevant data elements (e.g. using XML approaches) within the documentation collected aren't adopted. Instead, InfraDig errs towards capturing more than might be strictly necessary but which is easy to collect. While it has some resources to develop a 3D prototype, it doesn't have resources to really challenge or facilitate changing Operator practice, Data Engagement and learning.

Both initial Acorn Data collection and InfraDig's asset Data collection are revealed as necessarily large-scale, collective efforts. This respectively reflects the number of households across the UK and the vast amount of Design and As-built Documentation to be collected about a large infrastructure from various construction Contractors. In addition to pragmatism about data collection effort, the Data collected in both cases also seems to favour objective, consistently measurable data. Again, such data may be easier to collect and capture. However, differences in the underlying Phenomena can also be discerned as a potential influence. For example, much of the Data collected at InfraDig relate to physical characteristics of the infrastructure. These Data are relatively straightforward to measure, record and present, as they build on considerable consensus within Engineering communities. However, measuring likely life-time infrastructure costs (e.g. formulating Total Cost of Ownership models) and maintenance emerges as contested. Tacit construction (how to) experience is also difficult to capture and reflect in codified form. This is recognised and addressed instead by seeding the project with operator staff and planning for extensive handover activity to facilitate tacit Knowledge transfer.

My data collection was also pragmatic, reflecting the time I could spend in the field and emergent opportunities that arose in the field, e.g. to map Practitioner Groups at InfraDig. Given the social Phenomenon in view, this prompted me to carefully consider how to gain appropriate access (Van Maanen: 2011, Gummesson: 2000) and the nature of the Data to be collected (e.g. direct observation, meeting records, etc.). Various qualitative data analysis approaches were used to prompt and facilitate reflexive reflection. The research process' clarity and awareness about relating the Data collected to the

particular Phenomenon of interest and what Data analysis methods were therefore possible or desirable, was in contrast to relatively ad hoc related decisions observed in the cases studied.

Data evolution during exploratory Engagement

Considering my Data analysis reflexively highlighted how my Data evolved over various iterations of Data Engagement, reflecting greater progressive clarity, as outlined in section 7.5.4. During coding new codes emerged and related classifications were elaborated. These stabilised after several iterations, with further insights then emerging from different forms of analysis, reflection, writing, discussion and further reading. By contrast, among the cases, only at OtherCouncil did I observe a similar evolution and gradual stabilisation – of their deprivation index dimensions and related Data elements.

The elaboration of my coding classification is consistent with Tsoukas' (2005) processual definition of knowing, as the ability to make increasingly fine distinctions about a Phenomenon. The lack of Data evolution noted in the main cases seems to coincide with their relative lack of iterative Data Engagement. However, InfraDig also illustrates how Data stability may reflect a practitioner view that what is relevant, e.g. about maintenance, is settled and the related Knowledge mature. This highlights an inherent danger in equating stability with maturity, framing institutionalised Knowledge as satisfactory, which may act as a barrier to further evolution, insight and reframing.

Reflecting on my research also showed how Data's format and organisation can evolve, rather than just their 'content' or 'categories'. This emerged most clearly when I established a need to reorganise my Data collected along a timeline for each case. This was prompted by trying to undertake a new form of analysis, to establish how my Data related to the overall flow of events.

Relating this to OtherCouncil, a similar evolution of Data structure and 'content' can be discerned in their evolving dimensions of deprivation, eventually settling on a flexible index to encompass and address different users and context. OtherCouncil also provides an example of how such Data change and evolution

can be prompted by a change in Purpose. This prompts them to supplement or enrich the data, Changes in Purpose and use, and also seemed to prompt a search for related Data sources that overlapped, allowing a better combined Data representation of the Phenomenon of interest or might facilitate triangulation. In the InfraDig case this was seen in the KPI project, when it was established that existing project performance and timeline Data in the DW could be combined with asset quality Data, with the contract reference established as the common key. This would also allow relating progress in Data collection and quality to project timelines and delivery milestones. Discussion of common keys was also a feature of discussion about efforts to align Operator and InfraDig Data elements to facilitate integration.

7.6.3 Data as a Boundary Object or Infrastructure

This is most immediately visible at InfraDig, when considering the proposed 'virtual' infrastructure Data artefact. Earlier, design documentation was already identified as a Boundary Object or infrastructure (Star: 2010), spanning two primary boundaries. On the one hand, it facilitates *coordination* between InfraDig design engineers and various construction contract engineers. On the other hand, together with as-built documentation elements, it spans the boundary between both of these groups and the ultimate operators. It is a relatively unstructured form of Data and represents a Sensemaking Artefact in relation to the built infrastructure. The shift from presenting it as a set of physical documents or files, to providing a repository of linked electronic documents, represents both an evolution of its form or format, though also introduces more structure by relating the documentation using a standardised hierarchy. Given the level of standardisation and institutionalisation involved, particularly across the first boundary between design engineers and Contractors, this tends more towards Boundary infrastructure rather than a Boundary Object to facilitate peripheral engagement.

Less obviously, Data's role as a Boundary Artefact or infrastructure can also be discerned at GoCouncil. On the one hand, it spans a boundary between Acorn, as Data provider and subsequent users such as GoCouncil. This boundary

partly reflects the separation of the Acorn Data design and collection effort, again highly standardised due to its scale and specialisation. The boundary also reflects the repurposing by Marketing of Acorn Data intended for public policy Purposes. On the other hand, the simplified target group profiles or summaries also represent Boundary Artefacts. While there is little or no interaction between Acorn and GoCouncil, the target category summaries facilitate Engagement across the boundary between Marketing and other functional users. They prompt a discussion about targeting particular customers and the market viability of particular business model ideas. OtherCouncil's development of a deprivation index, another Boundary Artefact, also facilitates a discussion between different departments and potential users about deprivation. This discussion encompasses which dimensions are important in different contexts, how best to measure these and where related Data may be sourced.

On first reflection, this idea didn't seem relevant in relation to my research. Yet, on further reflection, many of the analytical outputs created and shared with participants also represent Boundary Artefacts between myself, as researcher, and practitioners being researched. They facilitated discussion and Engagement with research ideas and concepts, e.g. the CoP mapping of Groups at InfraDig and 2x2 question analysis at GoCouncil. They provided an opportunity for me to share and validate the sense I am making of their practitioner context too, to discuss their practice and to gain a better understanding.

Authorship and ownership

Considering Data as a Boundary Artefact or infrastructure and Sensemaking artefact brings into focus the different Practitioner Groups involved in creating, managing and using the Data. This in turn helps clarify which economies of meaning and concerns the Data seek to address and incorporate or encode.

In the case of InfraDig, the economy of meaning reflected is that of the engineering practitioner community. While this is well understood by both design and construction engineers, it is perhaps less well understood by those collecting and organising the Data, i.e. document controllers and the asset data

team. For Acorn Data, their 'authorship' is more remote or distant, reflecting economies of meaning related to public policy practitioners now opportunistically re-appropriated for a Marketing targeting Purpose. Indeed, Tanya's prior public policy role and Acorn experience is recognised as important to interpret these Data and select what is relevant for GoCouncil's purpose.

Reflecting on the authorship of my own Data revealed an intense sense of ownership and identification with my Data that evolved during my research. This may reflect the level of time and effort invested in collecting, creating, and Engaging with those Data to analyse and understand them. Reconsidering the cases in this light prompted a recognition of a similar sense of ownership of the Data within the InfraDig Asset Data team, though perhaps less intense than mine. Similarly, at GoCouncil, Marketing and Tanya in particular seem similarly invested in their target category summaries in a way they aren't in Acorn Data more generally. At InfraDig a flipside reaction from Operators might be at work. InfraDig's proposed Asset Data artefact's different structure may be seen as something 'foreign', being imposed and competing with 'their' existing Data structures and related ways of doing things, which they are invested in and feel 'ownership' of. This may partly explain the challenges to agree on Data requirements.

Directionality of Engagement facilitated

Considering Data's boundary nature, I also observed how the nature of Boundary (Artefact) Engagement varied within the cases. The interaction facilitated across certain boundaries was also sometimes predominantly in one direction rather than than bi-directional, *though this was not a feature of my research on Boundary Artefacts.*

As already noted in the case of Acorn Data, the Data are made available to GoCouncil as a product, together with manuals and some training available. There is no real interaction between GoCouncil and Acorn Data suppliers, which simply represents a one-way provision of Data. This contrasts with the summaries, which facilitate discussion between Marketing and the functional teams. Functional teams can also request more detailed Acorn Data from

Tanya and discussions also highlight additional Data that might be useful, e.g. as captured in the 2x2 question analysis.

In the case of the boundary between designers and construction engineers at InfraDig, the direction of design Data provision is mainly from designers to construction engineers, framed as part of a contractual requirement. A limited feedback loop across the boundary is noted though, in relation to as-built changes to the designs. The automated 'red-lining' mark-up tool being developed, will facilitiate near real-time updating of as-built changes onto designs, which may well facilitate far more Engagement. Immediate feedback from construction engineers to clarify and suggest changes, may allow more time and scope for designers to assess impacts, discuss alternative options with construction partners, etc. Such increased peripheral Engagement may well in turn facilitate enhanced understanding and learning by both groups.

Turning to the second boundary spanned by the asset Data, between InfraDig and operator maintenance practitioners, almost no interaction is noted beyond the discussion of asset Data requirements between asset Data specialists. Here the Data will simply be delivered to the Operators at the completion of the project for them to make sense of. Indeed, the time lag between collection and eventual use, given the closure of the project and loss of access to related staff, will preclude opportunities for further Boundary Engagement.

Boundary artefacts can anchor meaning

The considerable institutionalisation of design documentation noted at InfraDig has evolved over time between designers and construction engineers. It represents a joint enterprise and considerable investment in shared meaning. Changes in these Boundary Artefacts and related Data are likely to require co-evolution and negotiation. This may well constrain natural local Data evolution reflecting changes and improvements in local practice or Knowledge.

Much of this effort surfaces for participants as the need to *standardise* on various Data classifications, structure, etc. This is perceived as challenging and often described in terms of operator system Data integration. Relative influence

or power differences in relationships across boundaries also seem to play an important role in establishing and changing such standards and related economies of meaning. For instance, InfraDig has far greater scope (though still constrained) to influence and change the Data Boundary Artefact across its boundary with Contractors than it has with its ultimate clients, the Operators.

The greater the number of parties involved in such standardisation effort, the more challenging this seems to become. The challenge also seems greater for existing Operators compared to the new Operator. The new Operator doesn't yet have a similar, sunk investment in systems and processes and is happy to accept InfraDig's proposals. The standardisation effort and challenges noted seem similar to those noted by Ribes and Bowker (2009), to relate different Knowledge domains in order to facilitate cross-disciplinary working.

This may well be a factor in InfraDig's approach to collect and relate existing design and as-built documentation, rather than introduce a completely new data requirement and approach (e.g. introducing XML mark-up). While this approach eases the introduction of automation and eB for Data collection, it also entrenches existing practice and documentation approaches.

GoCouncil also has very little influence over changes to the Acorn Data and the supplier seems similarly constrained to make changes. It has much invested in the current Data collection approach and an interest in meeting broad client requirements, who may also value consistency and the existing coverage.

Both of these examples contrast with my research effort, where Data could evolve fairly fluidly without similar boundary anchoring during Data analysis; this reflects it being an individual effort without the need for similar coordination.

Focusing on producing Data can distract from the Phenomenon

At InfraDig many of those involved in Data collection and organisation are not aware of how it would be used, nor are they familiar with maintenance. Instead, they are purely focused on collecting, organising and producing a Data Artefact, as an end in itself. This contrasts with my *direct* Data collection and subsequent analysis and Helga's focus group effort. It highlights how, when there is

separation between Data production and use, Data producers may lose sight of the overall Purpose and objectives.

The scale of Data collection may be an important factor. Direct involvement from Data design, through collection and Engagement, was clearly more feasible in my research and in the GoCouncil focus group effort. Collecting and organising Acorn household Data across the UK, or all the design and as-built Documentation at InfraDig, necessitates considerable coordination and specialisation. This seems to correspond to a greater need to standardise Data definitions and classifications used to facilitate such coordination and consistency where multiple people or different groups are involved.

My complementary direct interaction and indirect Data Engagement and analysis also highlighted how simply being presented with the Data may preclude, or significantly reduce, the level of complementary direct Engagement with the Phenomenon of interest. It also highlighted the reduced opportunities for triangulation and how anchoring of attention on indirect Data Engagement might occur, especially where Data are closer to hand or all that are available.

7.6.4 Tools simultaneously enable and constrain

We have already noted how some Tools, such as eB at InfraDig, incorporate their own Data model and how this might frame thinking about a Phenomenon being addressed, in terms of Data dimensions and categories that may be relevant. Here we consider how Tool choice and use more generally enable and constrain both Data Engagement and exploiting related insight to generate benefits.

Different types of Tool facilitate different Engagement

When considered in terms of the Engagement they facilitate, three main groups of Tool emerge. First, there are various end-user Tools (e.g. Excel and SQL) that enable users to Engage very flexibly with a range of Data in a variety of ways. Second, there are functionally rich systems that incorporate or impose a specific Data Model and mediate Engagement with Data (e.g. eB or the proposed CRM system). In contrast to the end-user Tools, these systems seem

to tightly integrate or 'couple' Data and functional, algorithmic elements in a single, highly optimised Tool. Finally, there are systems or Tools primarily focused on Data presentation (e.g. the GIS tool and the 3D prototype).

The first group tends to be relatively straightforward to use, and under the control of the Data users in the cases studied. Excel, in particular, was widely used and understood, with an intuitive interface with which to Engage directly with Data. Relatively little Tool knowledge is required in order to start Engaging with Data using these Tools, although SQL does require some basic programming knowledge. The second group is far more complex, requiring considerable IT expertise to implement, configure, and then maintain and host. End-users have limited ability to affect significant changes and require specific knowledge about how to generate relevant reports and extract relevant Data from these systems. Alternatively, they are reliant on an IT specialist for assistance. This group of systems encapsulates a particular set of Data and algorithmic elements, grounded in previous experience of completing particular activities for particular Purposes or addressing particular Questions for a specific context such as construction.

The third group of Tools identified aims primarily to enhance Data *presentation* in order to enable improved user Engagement with the Data presented. In the case of GIS, this allows for the plotting of Data on a geographical map, to facilitate users making spatial connections and recognising related correlations and patterns. While these allow users a richer visual Engagement and exploration of the underlying Data, they also require specialist Tool or programme knowledge to use or develop.

NVivo and Prezi are the main Tools I used for my research Data Engagement. NVivo addressed a variety of different kinds of Data (e.g. audio and transcripts) and was especially useful to analyse and compare my interview Data. However, the techniques, forms of Data engagement and presentation were quite specific. NVivo also required specific training (a 2-day course and my acquiring a manual) to make the most of it. By contrast, Prezi was less functionally rich in terms of analysis but very intuitive and flexible to use after watching two or three

short online video tutorials. While NVivo clearly aims to enable Data Engagement and Analysis, Prezi is interesting in facilitating Data capture and Engagement as well as presentation, where I found NVivo to be less strong.

A predisposition to start with Tool or Systems selection

My initial Nvivo Tool choice was influenced by its being a default Tool recommended and available for analysing Qualitative Data Analysis at Cranfield, as well as empirical Qualitative papers reviewed – it seemed to be a de facto standard. This reflects a form of institutionalised thinking about appropriate Tools to use within research. However, as highlighted in Chapter 5, I was frustrated by its fairly narrow focus on coding and related analysis, in particular its inability to map Data on a timeline or analyse my coding progression over time. I was also frustrated by NVivo's clumsy and inflexible user interface. These shortcomings prompted me to use additional Tools.

At InfraDig, the eB and related design documentation IT solutions were selected very early on in the project, where they were seen as a prerequisite to the Data collection effort. Requirements were grounded in the prior experience of consultants involved in the selection process rather than ultimate maintenance or operator staff. Such 'IT system solution-centrism' was also noted at a wider industry level, when reviewing the COBie working group output (BIM: 2011). Its maturity roadmap focuses on how Tools deployed correspond to different levels of maturity, in terms of the Data and analytical functionality this enables, in order to better model the built infrastructure. A requirements-led, sequential systems selection and implementation process further reinforces the pre-disposition noted to *start* with selecting a Tool or IT system solution. This was particularly prominent at InfraDig.

At GoCouncil, this emerged when participants recognised a Data gap in relation to Data on existing customers and their service use. They immediately identified the need for a CRM system to address this, even though their related Data requirements are not yet clear and they have an incomplete view of existing customer service Data they hold. This contrasts with OtherCouncil, where existing systems, together with some end-user, general purpose Excel and SQL

Tools, are used. Considerable progress was made without the need for additional Tools. Only having achieved stability in their Data requirements do they consider using IT to develop some specific, automated routines or programs to optimise their Data processing.

Such Tool-centrism diverts participant attention from pragmatic Data Engagement and the Phenomenon, and instead focuses it on relatively abstract selection and implementation considerations and criteria. This increases the risk of selecting an inappropriate Tool or system and constraining the kind of Data Engagement and analysis possible.

Tools can frame Data Engagement and Phenomena

Another reason for highlighting the Tool starting point, is the way in which some Tools incorporate and represent Data about various Phenomena or dimensions of interest. This emerged particularly clearly in the InfraDig case, with the centrality of the eB Tool. This came into particular focus during the asset data KPI project.

The datafields and functionality available within eB represented the starting point for thinking about what KPIs and measures were available and could be selected. The development manager, with a broader IT perspective, sought to point out the possibilities of integrating Data from other systems, e.g. the project tracking system using the corporate DW. However, this requires coordination with other Groups and additional development cost, so is 'parked' for further consideration in due course, and attention returns to what is straightforward to achieve within the software that is familiar and readily available.

The eB system's inherent design focus on tracking Data and documentation quality and completeness, means lots of measures are available in this area. However, the outcomes they seek are predicated on achieving or shifting behaviours and mindsets of key project and Contractor staff (e.g. to recognise the importance of Data). No related measures are available though and this can only be triangulated or inferred, based on which contracts (and therefore staff) are not meeting Data quality outcomes, which requires additional Data to be

obtained and integrated from the project tracking system. Such factors led to a preponderance of lagging, indirect indicators being identified, which may hamper early intervention and more targeted approaches to particular individuals who might not be 'on board'.

This anchoring of thinking in what is available within the supporting application seems implicit and strong and may have been reinforced by the framing of the project as an IS systems project, with an IT Contractor as project manager. This further highlights the role of the application designer, as the application incorporates a 'snapshot' of the datasets and datafields *they* thought relevant at the design stage. However, their Purpose, thinking, view of the Phenomena they were focused on, and related configuration and design assumptions, are not readily available for reflection and review by InfraDig, given the third party nature of the software. Data Engagement was further hampered by the supplier's use of a proprietary database rather than a standard SQL database. This prompted the asset Data super-user to export Data extracts to a more flexible Tool such as Excel to facilitate easier, more intuitive exploration, manipulation and presentation of the asset Data.

As already highlighted, Tools also constrained my data analysis. NVivo's design reflects optimisation in relation to its original Purpose, and the approaches and practices it aimed to support. However, such original Purpose and context is not always obvious to a new user, unfamiliar with the Tool and its history, so the constraints remain hidden from view until the constraint is encountered. For me, these design constraints only emerged during use though, and represent learning to consider more carefully a Tool's original design objectives and optimisation in relation to what I have in mind.

7.7 Barriers to Engagement

A central idea reflected in the emergent explanatory framework, is that learning from Data occurs through Data Engagement. Having examined how Tools can both enable and constrain such Engagement, this section goes on to examine five additional Barriers to such Engagement that were observed across both cases. These seem, in part at least, to explain the missed opportunities noted,

particularly at InfraDig. This idea of Barriers to Data Engagement also represents the final refinement to the provisional explanatory framework to emerge from the case comparison. It is illustrated in Figure 7-6 by introducing the coloured bar across the Indirect Data Engagement Arrow.

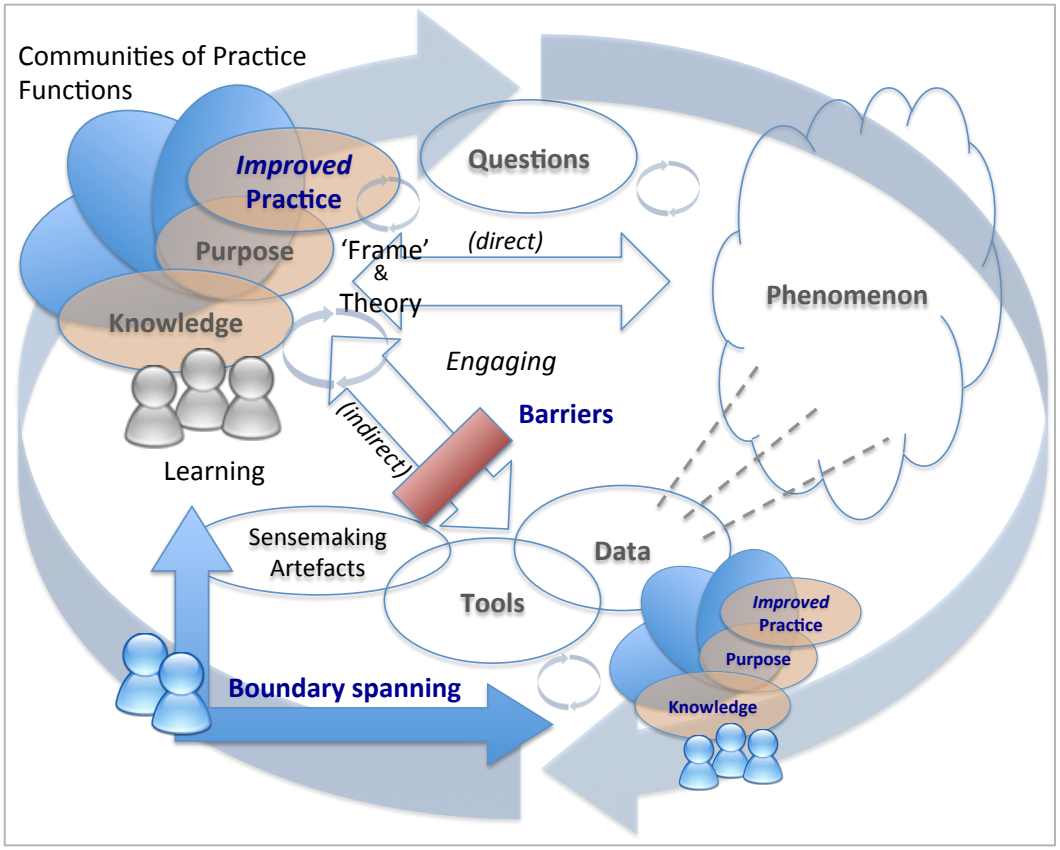


Figure 7-6 Barriers to Data Engagement as a framework refinement

The additional Barriers observed are considered in turn, although many seem to occur together, reinforcing each other. They seem to represent 'hurdles' for participants to effective Data Engagement and related learning.

7.7.1 Knowledge gaps introduce path dependencies

This Barrier revisits the question of domain Knowledge the initiative draws on, and related gaps, first touched on in section 7.4.2. In addition to wider impacts on framing and pursuing Data initiatives, these gaps were also observed to have a significant impact on Data Engagement.

In the case of GoCouncil, this was most clearly seen in the lack of Marketing and wider commercial Knowledge to make the most of Acorn Data. Improving these areas of domain Knowledge and experience was recognised as important by GoCouncil and reflected in their plans for business model training, although subsequently put on hold due to funding and time constraints. In the meantime, participants relied on those with more Knowledge, such as Tanya, who then rapidly can become a constraint or bottleneck. GoCouncil also highlighted how a lack of participant Knowledge about end-user Data tools (e.g. SQL) and the Data the council held hampered using and Engaging with the Data they have.

Turning to InfraDig, the area of Knowledge participants highlighted as most significant was the Operators' relative lack of Knowledge about using Data effectively in the maintenance context. Yet, as already highlighted, little was done to specifically address this or to afford opportunities for coaching and experimentation to aid related learning to bridge this gap. With the exception of one or two Boundary Spanners with prior maintenance experience, InfraDig project participants' Knowledge about maintenance or Operator contexts may also represent a gap. Similarly, there is no recognition of the lack of finance and accounting Knowledge.

InfraDig also relies heavily on the eB tool Knowledge of a particular individual in the Asset Data team. Again this represents a constraint for Data Engagement, especially given the quirks of eB, in terms of its proprietary query interface and Data structure. This person's experience of projects rather than maintenance also seems to influence the nature of the exploratory Data investigation undertaken, which is mostly project and data quality related.

Reflexively, this is mirrored in my need to master various new Knowledge domains in relation to my identified Phenomenon of interest (e.g, a literature review of Knowledge Management, Situated Learning, and Sensemaking domains); as well as research methods (e.g completing a research methods course, specific reading on how to undertake ethnographic research); and knowledge on how to use relevant Data Tools (e.g. NVivo for qualitative Data analysis). The path dependencies related to these are also partly reflected in

the overall research process and phasing, starting with a literature review, progressing to method and research design, before Engaging in Data collection and analysis.

These findings are consistent with ideas of path dependency highlighted by Cohen and Levinthal (1990) and their related concept of 'absorptive capacity', within research and development contexts. They are also consistent with the related need for a diversity of Knowledge and perspectives in teams to improve decision-making and innovation (Leonard & Swap: 2006). Both of these are in turn consistent with picking up on a variety of potentially relevant cues that might otherwise be missed (Tsoukas: 2005, Weick: 1995). Finally, the above examples also highlight how Knowledge gaps and path dependencies can relate to the Phenomenon as well as the methods, Tools and Data used.

7.7.2 Physical and temporal separation of Data collection and use

The next Barrier identified relates to how separation of Data collection and Data Engagement noted can hinder Data Engagement, at the same time diverting attention or distancing participants from the ultimate Phenomenon of interest.

At InfraDig, the physical and temporal separation of Data collection and use, already highlighted in 7.4.3 above, means that the ultimate Data users (i.e. operator maintenance teams) won't have access to the knowledgeable practitioners who created and organised the Data. They will only have access to the Data Artefacts themselves (i.e. design and as-built documentation). These, by definition, won't include the tacit design and construction Knowledge embodied in these Practitioner Groups, i.e. what Tsoukas (2005) describes as the inevitable precepts required to interpret codified Knowledge or Data. The Operator maintenance teams may not have access to the same Tools for managing and organising the Data collected, nor to the Asset Data Team who configured and structured the Data during the project. While the importance of tacit design and construction Knowledge is recognised elsewhere in the project, this doesn't address asset Data. The physical and temporal separation of maintenance and asset data Practitioner Groups at InfraDig also precludes the

opportunity for serendipitous Boundary Spanning interactions between these Groups.

The separation of Data collection and use or Engagement noted in relation to GoCouncil and Acorn has a similar impact. While Knowledge about Acorn Data collection and presentation is codified in training material, and some familiarity of it is reasonably widespread within the Marketing Practitioner Group, considerable investment is required on the part of less familiar users to fully understand and use Acorn Data effectively. This contrasts with the focus group activity undertaken by Helga, where they are more directly involved in both collecting the Data sought and its subsequent interpretation and use, and no similar Barrier was noted.

A similar continuity of involvement during my Data collection and analysis, certainly afforded an appreciation of tacit aspects. These aspects, sometimes called 'head-notes' (Schultze: 2000), could not be fully reflected in the Data or reflections captured, highlighting Data's inherent limitations. In common with Helga's focus group engagement with intermediaries, my research highlights the value of complementary direct Engagement with the Phenomenon, in addition to indirect Data Engagement. This also highlights how Data itself introduces an inherent distance and separation from the Phenomenon of interest through abstraction and interpretation, which is heightened when Data collection and Data use or Engagement are separated.

7.7.3 A lack of time, resources and space for exploration

Closely related to the *separation* of Practitioner Groups and activities, the time (and related mental 'space') to devote to Data Engagement, especially exploration, surfaced as a significant constraint or Barrier for participants across both cases (see Appendix D.4). This seemed particularly true for those participants for whom Data related activities were a part-time or secondary activity, i.e. not their main job.

Interestingly physical spaces were created in both cases: InfraDig's academy space, created with its software supplier to explore and play with new Tools and

functionality, and GoCouncil's Engine Room. These differed in that the GoCouncil space was specifically made easily accessible to its target audience to promote Engagement and use, while the InfraDig academy was at the software supplier's premises, which was not readily accessible for any of the key Practitioner Groups in view. Unsurprisingly, participants didn't use it and its main function seemed to be for new joiner induction sessions.

At GoCouncil, while the deputy CEO recognised the need for resources to be dedicated or 'ring-fenced' to work on business models, he also expressed a concern about doing Data analysis for its own sake, i.e. analysis which wasn't clearly *directed* and designed. Tanya and Albert try to point to the complementary value of more *exploratory* Data Engagement, using the example of serendipitously noting an interest in gardening while producing target summaries. This concern seems to be driven by the wider *urgency* to develop and launch new business propositions as soon as possible. Against this background and framing for activity, exploratory activity is seen as wasteful and to be avoided.

At InfraDig, the most striking such example noted was a sudden challenge and curtailment of innovation initiatives, when the project is identified as behind schedule, and therefore over budget. It was reflected in a loss of enthusiasm for Boundary Engagement with Operators in relation to asset Data. This also resulted in a shift from more exploratory focused Boundary Engagement, to the more specific and limited Engagement required to secure agreement on requirements and coordinate activity. This shift illustrates the vulnerability of Boundary Spanning activity, and how it can easily become seen as relatively less necessary when a practice area becomes resource (or time) constrained. Practitioner Groups seem to narrow their focus and prioritise their immediate, core or primary activities and objectives. Yet, to achieve InfraDig's wider objective for adoption and use of the Data being collected by Operators, such exploratory Boundary Spanning Engagement seems particularly important.

In contrast to these examples, which characterise investing time in clearly directed Engagement as more productive than exploratory activities, my

research experience points to the considerable value of exploratory Engagement activity. Sometimes this occurred as a complement to more directed analysis and at other times was a precursor to it. My research experience also highlighted how research required considerably more time for data analysis and reflection than was ever available to me or possible in my prior consulting experience.

The above illustrates how both exploratory Engagement and Boundary Spanning activity often require additional 'space' and time over and above core or primary and immediate practitioner activities and objectives. At the same time it highlights their sensitivity and susceptibility in the face of constraints. These observations are consistent with findings in relation to incremental learning capabilities, which stress the importance of slack resources (Salge & Vera: 2013). They also recall Weick's (1995) observation of a focal narrowing that occurs under extreme pressure, which may be relevant for missing potentially relevant cues when undertaking Data Engagement under pressure.

7.7.4 Practice Edges and Boundaries hardening into Borders

Many of the earlier Barriers reflect or stem from the Boundary-Spanning nature of Data initiatives first highlighted in 7.4.1 above. This Barrier describes the observation that some practice edges identified vary in how high a hurdle they introduce or represent for Boundary Spanning activity. This seems to coincide with the degree to which boundaries and edges between different Practitioner Groups are situated within common organisational units or span different units and organisations. When these occur within a common unit they might perhaps be characterised as *soft* or *permeable* boundaries and edges. To the extent they span department, functional or organisational boundaries though, they might be characterised as progressively *harder* or *less permeable*, gradually crystallising into more formal *borders*. In both cases Practitioner Group delineation seemed to coincide with functional organisation units, especially for well established and common specialised practice areas such as Finance, HR and IT. In addition to these, engineering design within InfraDig, and Marketing within GoCouncil, were also identified in this category.

As would be expected, different Practice Groups clearly showed distinct identities. These identities also transferred to or became associated with the related Data Artefacts and Tools they Engaged with. For example, IT identified with corporate systems while engineering identified with their own end-user developed Tools and specific eB system and Data in the InfraDig case. Similarly, the virtual infrastructure Data initiative and Artefact became closely associated with design engineering. However, it was various Practitioner Groups' different 'economies of meaning' (Wenger: 1998) that surfaced as particularly pertinent for Data Engagement. Terms and frames of reference used reflected the dominant or leading Practitioner Group involved. In the case of InfraDig it reflected design engineering usages and data structure, while in the case of GoCouncil it reflected Marketing and Acorn usages and terminology. Other Practitioner Groups were expected to Engage using the dominant Group's terminology and concepts, which are also inherently reflected in their Data Artefacts. This aspect of the Barrier is tightly interrelated with the Knowledge gaps and the related Barrier already highlighted in 7.7.1 above.

To the extent that Practice Group boundaries or edges coincided, or were reinforced by organisational functional units or delineation, this seemed to exacerbate these effects. Firstly, it formally institutionalised these differences and identities, reflected in physically separate work locations. For example, at InfraDig, IT was seen as a quite separate, distinct and distant group that merely hosted systems and provided basic IT infrastructure and office applications (e.g. email). They were also not rated highly by the Asset Data Team, given their limited understanding of asset Data applications and issues.

Secondly, these units had separate budgets, cost centres and formal objectives, which often reflected local rather than organisational priorities. Again this was highlighted at InfraDig when IT's initial enthusiasm to assist the Asset Data Team is overtaken by the pressing need to address a serious email systems issue, which required their primary focus. It was even more pronounced in relation to getting Contractors to address documentation Data shortcomings, especially at contract close. This second aspect highlights a direct connection

between these harder 'borders' and the susceptibility of coordination efforts across such boundaries or edges in scenarios of resource scarcity or urgency for either or both units. Such tensions were less pronounced or visible at GoCouncil, which may have reflected the initiative's strategic importance and related senior management support and resourcing.

Thirdly, differences in 'economies of meaning' and related tensions noted, were more pronounced between arms-length parties from different organisations. This was most noticeable between InfraDig and Operators in trying to agree on Data requirements, where different classifications and structuring were used for Data between different Operators, as well as InfraDig. These different economies of meaning were reinforced and reified by different IT systems or Tools used, which represented considerable sunk investment for them.

On the other hand, the effects of 'harder' boundaries or edges, were ameliorated somewhat by shared economies of meaning, for instance, where practitioners from different Practitioner Groups in different units or organisations were also members of a wider professional community, e.g. engineering. By contrast, different Practice Groups in the same organisation from different wider practice areas, which shared fewer concepts and economies of meaning, e.g. finance, seemed further apart, even though they might be expected to share greater goal alignment. This was reflected in the complete lack of involvement of Finance. It was underlined by an interview with the larger infrastructure organisation. They cited duplicate initiatives by their Asset Data and Finance Teams, generating different asset data classification systems, which was only identified during an external review.

7.7.5 Constraints on Iterations

Given the time required for iterations of exploratory Engagement and refinement, time and urgency related constraints seem particularly pertinent as barriers to such iterative Engagement and learning.

The GoCouncil deputy CEO's challenge of how much Data analysis is enough mirrors a similar research consideration. This question seems closely linked to

the characteristic of confidence. In the context of my research, this manifested as confidence first in my approach and research design, and then gradually in my findings during Data analysis and reflection. This was further strengthened or 'forged' through supervisory discussion and challenge, peer review of papers, colloquia and the like.

In the context of GoCouncil, this was evident in securing the confidence and backing of senior management and the board for the proposals. By contrast, OtherCouncil pointed to the need for the Data to provide transparent grounding for their service proposals to citizens, which seems a more defensive, evidential role. However, this too can be characterised as confidence building within the citizenry, as to the perceived fairness of the proposed service changes based on the supporting Data provided and the readiness to be transparent.

At InfraDig, no iterations were noted, or refinement. Here confidence to act and proceed seems to be grounded in the prior experience of maintenance of some of the key boundary-spanners within the engineering team. Iterations of learning only seem likely or possible across projects in this scenario, over very long timeframes (given the nature of these projects), and even then may still miss subsequent learning through actual maintenance experience.

These observations are consistent with Weick's (1995) ideas of enactment based on partial information, as well as ideas about satisficing (Davis & Olson: 1984). Weick's related concerns about equivocating in the face of ambiguity may also be at the root of the InfraDig IM strategy procrastination.

7.8 A SUMMARY OF COMPARATIVE CASE OBSERVATIONS

The reflexive comparison of cases outlined above both illustrated elements of the provisional explanatory framework introduced and generated additional observations and refinements. These additional insights emerged through the significant reflection it facilitated about the differences and similarities between the cases, my research and additional interviews with two additional organisations (OtherCouncil and InfraBig). For ease of reference, these comparative observations have been summarised in Table 7-2 below. This serves to highlight which explanatory model elements draw most richly from which cases, and to what extent observations and findings draw on a particular case, commonalities across the cases, and from reflexive comparison with my own research.

Table 7-2 Matrix Summary of Observations by Theme and Case (reflexively including my research as a third case)

Framework Element Observation	GoCouncil	InfraDig	My Research
Clarity of Purpose	New improved service propositions informed by customer insights from Acorn Data	**Asset Data initiative:** Improve maintenance practice with better Data **Information Management Strategy:** Make ALL InfraDig Data easily available	Improve understanding of the social reality of Data initiatives pursuing insight
	Identified the need to also consider market insights and other data	Some lack of clarity and more purposes identified - Audit trail requirement (operations) - Demonstrate good BIM practice (CEO)	Gradual clarification of Phenomenon in view and clearer articulation
Inquiry and benefit logic left implicit	Priority household groups as wealthy Household unit of analysis Intermediary customers also relevant	Data use and benefit realisation not articulated Data produced will be used and benefits will flow	Formal research design Relates Data to Phenomenon Ongoing panel review and refinement
Role of Questions	Initial board questions not kept under review Workshops - they clarify and draw attention Abundance - posing a prioritisation challenge	Absence. Not question led, in spite of BIM illustrative questions Requirements as well understood/settled	Research questions central to inquiry and presentation of research Evolve and are refined
Practitioners shape initiatives			
Boundary spanning nature	Marketing (led) and functional departments Expressly framed as council-wide	Various specialised groups involved in data design, collection, management and use. Engineering led. Initiative spans organisational boundaries	Spans IS, Learning and Knowledge Mgt. disciplines Research represents a form of peripheral participation/ boundary engagement in cases

Framework Element Observation	GoCouncil	InfraDig	My Research
Role of Group Domain Knowledge	Marketing domain knowledge Acorn Data knowledge	Various teams bring specialised knowledge (infrastructure, asset data, tools/systems)	IS discipline Panel add learning and Marketing perspectives
	Commercial and business model knowledge gap emerges as important	Maintenance, finance and sustainability missing Boundary spanner reliance (maintenance)	New to Learning and Knowledge Mgt. domains New to Ethnographic research methods
Separation of Data collection and use	Acorn Data obtained from specialist provider Marketing summaries for functions to use Integrated for intermediary focus groups	InfraDig design and collect Data with Contractors Operators will use the Data after handover (physical and temporal separation)	Integrated data collection and use Individual rather than collective effort
Challenges aligning practitioner groups	Peer groups - limited influence on Acorn use Challenge of how much insight is enough Use of Board platform to argue for benefits and challenge business models proposed	Engineering has some influence over Contractors CEO support given BIM Less influence over operators to use data Use boundary spanning, contracts, KPIs, Quality Assurance, management forums	Aligning research and practice agendas (e.g. emergent workshop and CoP opportunities)
Iterative Engagement and progressive clarity			
Data Engagement as central	Data led - Acorn Data Mainly by Tanya/Marketing team Some function engagement in workshops Complementary direct engagement for intermediary focus groups	Data led - Asset Data Focused on data collection and organisation Limited data engagement focused on data quality No engagement with maintenance Phenomenon Potential data integration hurdles for operators	Phenomenon led Various forms of data engagement (as per chapter 5)
Practice frames Engagement	Marketing - customer and market research Qualitative and quantitative approaches used	Engineering – waterfall project approach and standard construction documentation	Discipline norms as to method and presentation
Iterations of clarification	Some – target group assumptions and intermediary identification informal rather than planned stages/steps	Opportunities missed to test and refine data Little progressive clarity noted and no iterations Several assumptions as to use remain untested	Several iterations of refinement and clarification covering Phenomenon, research questions methods, data, findings and contributions

Framework Element Observation	GoCouncil	InfraDig	My Research
Data, Tools and Artefacts			
Data and Tools used	Acorn Data, Intermediary focus group Data, Target Household Summaries Excel, GIS Mapping Tool	Virtual infrastructure data artefact eB asset data management system, Excel, 3D visualisation prototype	Interview and meeting data (various, per ch. 3) Question mapping, Community Mapping NVivo, Prezi
Data as socially constructed snapshot of Phenomena	Acorn – for various public policy Purposes Tanya translates and filters for functions	System reified view of relevant data (eB) Use of institutionalised documentation structure	Clarity emerges through reflexiveness about data Active in collecting and creating (e.g. CoP maps)
	Pragmatic household unit of measure versus intermediary, requirement for new data	Standardisation challenge with operators Challenge to capture tacit aspects Seeing data need as settled may constrain learning and data evolution (e.g. use of XML)	Evolution of coding, especially initially
Data as boundary artefact	Acorn data from provider to GoCouncil Summaries from Marketing to functions	Primarily between InfraDig and Operators, though also between design engineers and Contractors	CoP mappings between myself and participants
		Data collection distracts from the Phenomenon Scale requires separating data collection and use	
Tools enable and constrain	Need for a CRM system identified before current Data and needs understood	Selected very early to facilitate Data collection (based on consultant and prior experience) less focus on requirements for maintenance use	
	Constrained by lack of tool knowledge (e.g. Access/SQL) GIS tools aid visual data engagement	Data tightly coupled within eB - proprietary Fields available anchor and bound KPI thinking Enables data collection – Contractor coordination but hampers access to data – extracted to Excel	NVivo constrains timeline view - Prezi used instead NVivo required more training investment effort Prezi – less mediated & visual data engagement Inherent design optimisation for original purpose

Framework Element Observation	GoCouncil	InfraDig	My Research
Barriers to Engagement			
Knowledge gaps as path dependencies	Lack of Acorn, Marketing and commercial knowledge in functional departments Reliance on Tanya for Acorn a constraint	Lack of maintenance knowledge – user context Dependency on boundary spanners/consultants Dependency on eB super-user (and supplier)	Need to master Learning/Knowledge Mgt. fields Mastering ethnography research approaches and tools
Physical and temporal separation	Little ability to ask or influence data provider	No operator access to project staff post handover No refinement and testing interaction planned	Continuity of involvement (collection and use)
Time, resource and space constraints	Not a full time job, so limited attention Discouraged from exploratory activity	Curtailment of innovation activity when behind impacting on boundary activity with operators	Considerable exploratory activity Even so, limits on analysis - cannot be exhaustive
Boundaries hardening into borders	Functional and practitioner group boundaries often coincide within organisations, institutionalising identities	Functional and practitioner group boundaries/ identities often coincide within organisations, Non-alignment of incentives (e.g. Contractors) Different groups excluded (e.g. Finance)	
Constraints on interations	None noted - urgency to launch new or revised propositions	None noted, focus on data collection rather than refinement and learning	Progressive refinement and reflection with cycles of data engagement, panel and colloquium discussion, as well as further reading and writing

7.9 A REFINED EXPLANATORY MODEL

A more complete picture of the phenomenon emerges from the case comparison than the provisional explanatory framework presented at the beginning of the chapter. This enriched view reflects some additional elements and the highly iterative, complex 'dance' between interdependent elements that emerges more clearly. This is illustrated in Figure 7-7, presenting a refined explanatory framework, incorporating several new elements. These are overlaid on the earlier framework, and highlighted in colour to differentiate them.

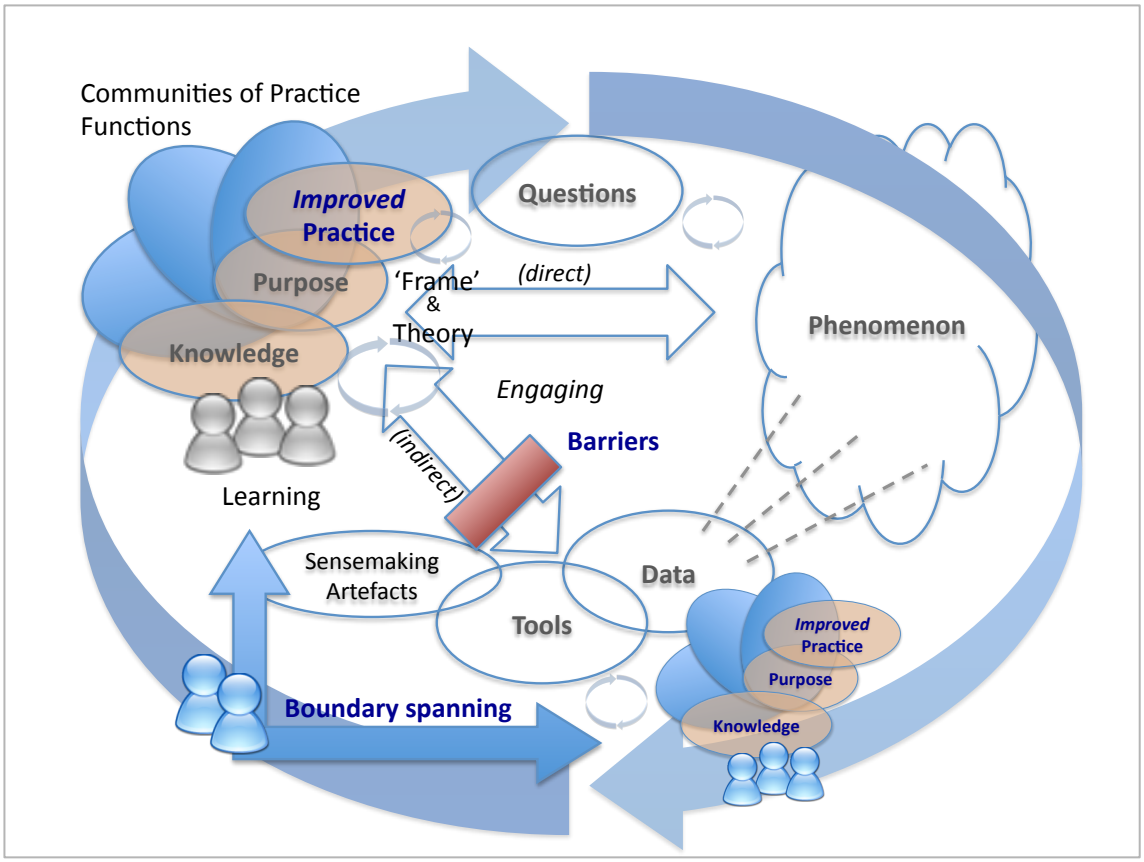


Figure 7-7 Refined framework for generating insight from data

This refined and more comprehensive framework is used as a reference point for the discussion, related arguments and contributions outlined in the next chapter.

8 DISCUSSION

Prior chapters have presented descriptive accounts and observations in relation to both cases, as well as a comparison of them with my own research experience. Taken together with this discussion, these seek to address the main research question posed: **How do participants in exploratory data initiatives collectively use data to pursue insight?**

This chapter focuses on the main theoretical contributions and implications that flow from the improved understanding of Data initiatives, as reflected in the explanatory framework presented in the previous chapter. These can be summarised as follows:

- A case to reframe our existing IS theory for data initiatives, as summarised in the MSS model, suggesting various enhancements to:
    - recognise and address the complex and interdependent inquiry and benefit realisation logics or theories being pursued,
    - reposition MSS Use in terms of these logics, arguing instead for the central importance of Data Engagement,
    - introduce Data as an explicit and core construct, distinguishing it from algorithmic elements to better theorise MSS artefacts, and to
    - more explicitly address specific practitioner contexts involved, related Knowledge domains, alignment and coordinatition.
- Arguing for the value of adopting a situated learning, CoP lens for researching data initiatives, at the same time identifying a shortcoming in the framework.

The first set of contributions relates directly to the research question and phenomenon in view, while the second contribution relates to using the CoP framework, representing a contribution related to method. These are discussed in turn below. The last section summarises the proposed contributions in tabular form, linking these to the main field observations summarised in the previous chapter, and gaps in existing theory identified in the chapter 2 literature review.

Note: Capitals are used when referring to framework or theory constructs.

8.1 Reframing and improving our IS models for data initiatives

The clearer understanding of Data initiatives, reflected in the explanatory framework, is related to the MSS model outlined earlier. In doing so, I am seeking to build on the existing Knowledge the model summarises, and at the same time to make clear the contribution represented by the improved understanding that has emerged from the research.

The Research findings provide some evidence for the importance of several Tool related constructs and a number of relationships posited within the MSS model. However, doing so is not the focus of my research, nor for this discussion and proposed contribution. Instead, a number of missing model constructs, missing construct dimensions, and construct definition clarifications emerge from the case comparison observations and explanatory model presented in Chapter 7. This section extends the earlier critique of the MSS model set out in the literature review (see section 2.2.3).

The discussion of these proposed MSS enhancements is grouped into different, though interrelated strands. The first strand is concerned with the logics underpinning such projects and how these prompt a reframing of the MSS model, primarily by revisiting what is termed the Problem Space and related constructs. The use of IT artefacts is then repositioned or reframed in relation to these logics. Another strand relates to giving prominence to Data as a separate construct within the model, disentangling it from the wider MSS artefact, in particular distinguishing it from the algorithmic components of an MSS, in order to better theorise both constructs in the context of data initiatives. The final strand addresses practitioner groups involved in such initiatives, to account for the specific domain Knowledge they contribute to the initiative, as well as their alignment and coordination. These strands identify several missing constructs that should be included and propose various refinements to existing MSS constructs.

These new and modified MSS constructs are explained and positioned visually in relation to the explanatory framework presented at the end of Chapter 7. Framework 'elements' are referred to in order to distinguish them from MSS 'constructs' during this discussion for greater clarity.

8.1.1 Explicitly address inquiry and benefit realisation logics at work

The broad argument put forward in this section is grounded in the idea that logics provide a frame or paradigm for the activity of pursuing insight and related benefits that will flow from them, i.e. a theory of action. Weick (1995) argues for their importance for participant sensemaking and related enactment of meaning. Quoting earlier work, he argues: "Theories of action are for organizations, what cognitive structures are for individuals" (Weick: 1995: p.121) and that these are important for participants and related groups as they build Knowledge when responding to situations encountered.

Weick (1995) links these ideas to various concepts such as cognitive, causal maps, Knowledge structures and mental models, arguing they all provide a frame for sensemaking, alerting participants to relevant cues and linking these to a repertoire of potential responses. Drawing on ideas from Argyris and Schön (1978, 1974), he also cautions that espoused theories may not match actual theories-in-use which produce enactive sensemaking.

As highlighted in the introduction and again in the initial critique of the MSS model in section 2.2.3, many Data initiatives and related investment in Tools seem to be predicated on relatively straightforward underpinning logics which could broadly be summarised as follows:

More Data + **Better** Tools = **More** Insight

and

Users + **Improved** MSS Tool (implementation) = **Improved** Quality Decisions

The latter logic reflects the MSS model's central focus on implementing better Tools to improve decision quality outcomes. Indeed, in both cases, participants reached for Data and Tools as a starting point, perhaps reflecting these

pervasive logics. This highlights the potential conflation of Tools and Data noted in such logic, especially evident in the InfraDig case where a focus on the eB system dominates. In GoCouncil's case, this is evident in their perceived need for a CRM solution, rather than using their existing systems and Data as the starting point for better customer insight.

However, the initiatives are more fundamentally underpinned by theories of how Data will shed light on a Phenomenon, as well as theories about how related insights might inform and improve related practices and activities. This points to actual theories-in-use which are often left implicit. In the case of GoCouncil, Acorn household Data are intended to shed light on customer service preferences, to inform the prioritising and designing of new and improved services. In the case of InfraDig, collecting and making better asset Data available to Operators is to yield insights with which to improve maintenance practices, in order to reduce lifetime maintenance costs.

Such implicit theories-in-use reveal two distinct, interlocking logics, one related to inquiry about a Phenomenon and the other related to a theory of benefit realisation, based on anticipated insights that might flow from the inquiry. This aligns to distinctions Argyris and Schön (1978) draw between theories or logics aimed at understanding versus those aimed at application. Based on this view, the inquiry logic also serves to distinguish *exploratory* Data initiatives from other Data initiatives that seek simply to apply existing theory, i.e. where the inquiry logic element may not be relevant, or where this element represents a decidedly secondary focus. These two logics or theories of action are examined in turn, as well as how they might evolve, before considering the implications for reframing and improving the MSS model with exploratory Data projects in mind.

Inquiry logic

The likely importance of an underpinning inquiry logic emerged early in the research process, given recursive similarities between the cases pursuing insight from Data and my own research seeking to do the same. In particular, this pointed to the relevance of research logic and concepts (Blaikie: 2010), and indeed proved a rich vein of reflexive insight.

While particular inquiry logics may vary in detail, depending on the particular context, the initial, simplified explanatory framework for the cases examined highlights how Data project inquiry logics are all likely to involve implicit or explicit Questions about a Phenomenon of interest, which they seek to answer by Engaging with related Data in order to improve practitioner Knowledge about the Phenomenon in view. Therefore, such logics will always involve the following key elements shown diagrammatically in Figure 8-1.

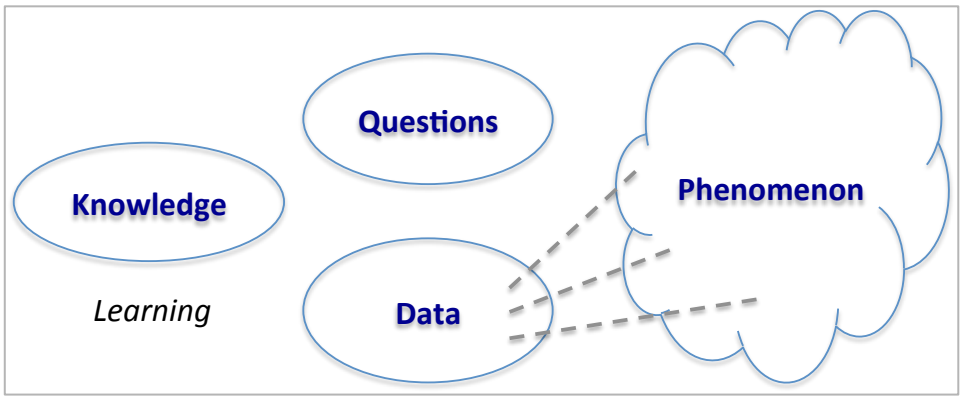


Figure 8-1 Key elements of an inquiry logic (practitioner group level)

Central to such logic is how well any Data to be analysed represents the Phenomenon for a particular Question posed. For instance, in the case of GoCouncil, Helga identifies that Acorn Data are not relevant if the customer is likely to be the intermediary rather than the household, or when OtherCouncil is interested in individual householders and needs to supplement the Acorn household Data with service data to address this. These examples point to the value of research concepts of measurement and validity as important considerations for Data projects (Easterby-Smith et al.: 2008).

The case comparison illustrates how practitioner Data Engagement is key to pursuing insight from Data. Particularly in the case of InfraDig, it also shows how not all participants were involved in Data Engagement and how several opportunities for Engaging with Data were missed. This can clearly be influenced by various individual and social factors. However, an initiative's underpinning inquiry logic also needs to anticipate and account for possible, indeed likely, *direct* practitioner Engagement with a Phenomenon of interest.

This is depicted in Figure 8-2. Helga at GoCouncil provides an example of this when she approaches intermediaries directly, to understand their needs, rather than seeking out existing Data about them. In this case, the related information and Knowledge gained remains tacit, held by Helga and other workshop participants, rather than being crystallised in formal Data collected.

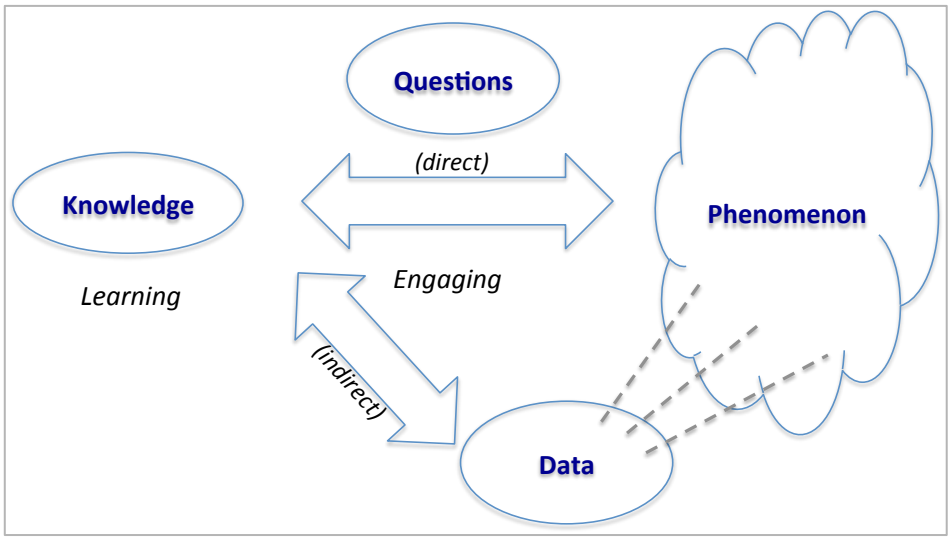


Figure 8-2 Key elements of an inquiry logic – direct and indirect engagement

On the one hand, such direct Engagement with a Phenomenon could act as a substitute for Data Engagement, perhaps undermining the rationale for the Data initiative. Direct Engagement can also lead practitioners to privilege their direct experience and Knowledge during Data selection and Engagement, although some privileging of prior Knowledge or its framing inquiry may be unavoidable in any event. However, on the other hand, direct Engagement with the Phenomenon can also provide a complementary perspective to that offered by Data Engagement. In this case, it may offer opportunities for triangulation, as well as validation. This seems akin to mixed methods research, combining action research elements with other more Data-led methods. Such validation may in turn enhance trust in such Data and related analysis, or prompt Data quality improvements, especially where practitioners are not directly involved in Data collection. This accords with earlier work highlighting the likely importance of trust for information orientation and information or Data management (Marchand et al.: 2001).

The degree of direct Engagement with a Phenomenon of interest may be influenced by the nature and scale of the Phenomenon itself, in terms of the feasibility of direct Engagement or observation. For example, quantum Phenomena or distant galaxies preclude direct human observation and Engagement, falling outside our cognitive capacity or range, necessitating an indirect, data-centric approach. This is evident in GoCouncil's use of secondary Acorn household census Data where it wasn't feasible for them to Engage with all households in their district to collect relevant Data. This example highlights the degree to which logistics and measurement challenges may also vary depending on the nature of particular Phenomenon, i.e. assessing or measuring and recording particular states in relation to dimensions of interest about a Phenomenon.

In addition to questions of scale, inherent measurement challenges are also likely to arise when the inquiry relates to shifting social Phenomena compared to physical Phenomena, subject to fixed permament laws. Social Phenomena are therefore likely to prompt more longitudinal data requirements and analysis, and result in more provisional, context bound insights. Related analysis will also involve different validity considerations, often more complex and potentially contested.

Benefit realisation logic

In contrast to the inquiry logics underpinning Data projects, the importance of a clear, validated benefit realisation logic emerged during InfraDig fieldwork. This draws on work by Ward and Daniel (2012) within project and programme management, which seeks to understand the benefit objectives, related dependencies and assumptions as a starting point. In essence, this seeks to clarify a particular kind of theory-of-action. The InfraDig case highlighted the following benefit realisation theory being adopted: that if asset Data were collected and provided to Operators, they would engage with the Data to pursue and act on related insights about the assets to improve their maintenance practice. This in essence reflects the following logic, representing a slight variation on that identified earlier:

(providing) **More** Data = **More** Insight

What emerged during fieldwork was a series of important dependencies, and indeed impediments or Barriers, which might need to be addressed in order to encourage asset Data Engagement and to realise related benefits through practice improvements. Therefore InfraDig requires a more realistic logic or theory of how they will realise envisaged benefits, which recognises these dependencies.

The dependency on Operators Engaging with the Data collected to realise related benefits can also be argued to be part of the inquiry logic. This highlights the overlaps and interdependencies between the two logics at work. The distinction is clearer when considering GoCouncil. Here, insights did indeed emerge for those Engaging with the Data, e.g. Tanya in relation to the potential for new value added gardening services. However, related benefits were contingent on further investigation and action by the functional managers and teams they were working with. This situation seems very similar to that of Data analysts working with managers responsible for acting on insights provided to them (Wang & Wang: 2008). These examples serve to illustrate how such insights don't always feed into defined decision-making processes and may inform wider, less structured or bounded activities. They are instead characterised by related, embedded implicit decision-making.

The InfraDig case also provides an example of where the absence of a clear and compelling benefit logic leads to a lack of clear focus and procrastination. In the case of the wider information management strategy, no progess is made where there is no clear business imperative to devote time and resources to it. Instead, it seems to represent a vague, good practice goal to Donald for the organisation, reflected in recurring comments to lead by example by adopting good IMP.

My case observations support Weick's (1995) notion of how such benefit realisation logics represent a fundamental overarching context, providing an overall Purpose and frame of reference for any related inquiry undertaken. It sets expectations as to envisaged ultimate benefits, which in turn helps

prioritise lines of inquiry, and set an appropriate level of resources to devote to the effort and related constraints. This complements the earlier inquiry logic elements to include the following elements, as illustrated in Figure 8-3. It also points to their importance in achieving alignment of participants, argued by Wenger (1998) to be an important architectural element to promote emergent learning.

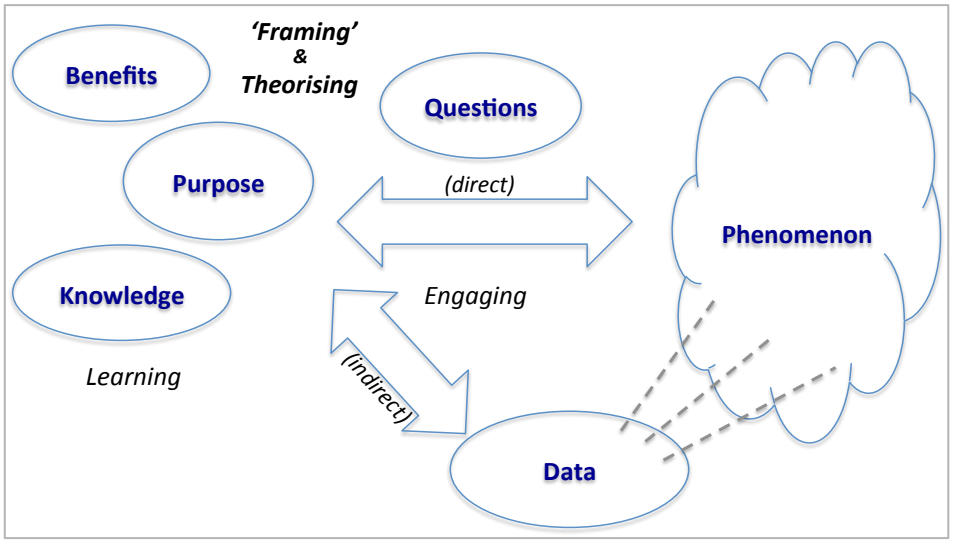


Figure 8-3 Benefits logic elements that frame the inquiry logic

Traditionally within projects, this logic is sometimes captured formally in a business case and assessed in terms of likely costs and benefits (Ward & Daniel: 2012). However, given the emergent nature of exploratory Data initiatives, it is firstly difficult to identify likely insights that may emerge, and secondly, to assess their potential value. Therefore, a research proposal approach (Marchand & Peppard: 2013) may be more appropriate, possibly combined with benefit option analysis for possible or likely areas of benefit.

In neither case are these logics made explicit. Instead, they are left largely implicit in the minds of key managers or participants leading the initiatives. Nor was any mechanism to review and refine these evident. However, greater clarity of understanding and grasp of the logic itself, of how benefits may be realised, would clearly be useful to surface key dependencies and assumptions. This facilitates benefit assessment and managing logical elements to improve the

likelihood of realising envisaged or possible benefits (Ward & Daniel: 2012), or to help determine the level of effort and investment warranted, e.g. for more tenuous benefits.

Realising or enacting such logics to crystallise benefits is clearly dependent on various organisational, social and individual factors, e.g. resources available, manager trust or confidence in analysts, and the ready availability of Data (Ward & Daniel: 2012, Davenport: 2014, Marchand et al.: 2001). This may serve to link elements of the logic to particular Practitioner Groups that may be important, which is addressed in due course.

Iterative clarification, alignment and validation

A clear grasp of the inquiry and benefit realisation logics being used by particular initiatives also brings into focus which elements and assumptions need further clarification, validation and testing. This aligns to what Weick (1995) describes as theories-of-action being constantly revised and refined through experience of applying them to select appropriate responses to recognised cues or stimuli. This was highlighted most clearly by my own research experience, characterised by several iterations of progressive clarification and specificity. In my research, this related to the Phenomenon in view, research Questions and how Data analysis might lead to insights, as well as Data's inherent limitations compared to direct Engagement with the Phenomenon.

As for the cases observed, the fact that these logics were left largely implicit meant that assumptions were only identified and challenged some way into the project. This can trip up initiatives in terms of reworking and expectations of progress, as in the case of GoCouncil, undermining confidence. More fundamentally, it can undermine the likely realisation of anticipated benefits, for example when at InfraDig it is assumed that Operators will use the Data provided to enhance their maintenance practice, representing a key assumption. This may well not happen given systems' incompatabilities, relative lack of practitioner Engagement planned during Data handover, or Data migration and integration planning.

The inherently iterative process observed during my research experience might also occur over the course of a series of related inquiries about a Phenomenon over time, rather than in a single initiative. This seems likely to lead to a gradual increase in domain Knowledge about a Phenomenon, with less frequent radical breakthroughs, as is common in research disciplines. Indeed the parallels with research seem profound. Learning across projects is also recognised as a specific branch of project based learning (Easterby-Smith & Lyles: 2003), which may prove useful for further theorising such contexts.

However, in addition to the obvious ongoing inquiry refinement, it is important to recognise that such iterations of inquiry may lead to double-loop learning or second order reframing (Argyris & Schön: 1978) of the wider Purpose and benefits that are likely to arise. Such progression and refinement also implies some sequential dependency between achieving understanding before moving to application and practice improvements acting on such improved Knowledge. This seems very similar to Cohen and Levinthal's (1990) notion of a path dependency, as part of their wider organisation level construct of absorptive capacity. Such iterations seem especially likely for exploratory initiatives where the Phenomenon is not well understood and the related domain Knowledge required is immature or tentative. Again, this seems typical of a traditional research and development approach.

However, we can also envisage practitioners simply acting, experimenting, testing or intervening in situations based on hunches they have for what may yield envisaged benefits or improvements. These actions may in turn lead to *post hoc* reflection, rationalisation and improvements in their understanding. In this regard, such approaches may benefit from adopting approaches and techniques from more recent engaged research approaches, such as action research (Eden & Huxham: 2002) and design ethnography (Baskerville & Myers: 2015).

Identifying the need for such iterations, and assessing the scale of ambition inherent in different inquiry and benefit realisation elements, are both important aspects of planning for such initiatives. Variations in iterations and ambition will

clearly impact on the likely scope, resources, costs and timescales required for such projects. For example, getting to the moon with untried technology required far more resources, experimentation, testing and time than constructing a bridge with tried and tested approaches. Periodically reassessing related planning assumptions is likely to be important for more exploratory and ambitious projects. For example, where several lines of inquiry are being pursued, regular progress assessment may help focus and prioritise limited resources or time on more promising lines and stop the pursuit of others.

As highlighted in the cases observed, these aspects may well crystallise various conflicting interests. The related power and resources of different Practitioner Groups involved or impacted on are likely to be an important consideration, as well as related powerful actors within and across such groups. Indeed, Weick (1995) highlights that strategic ambiguity can allow the perception of participant consensus to persist, when there isn't actual agreement. So, making the underlying inquiry and benefit logics explicit, is likely to reduce the scope for ambiguity, prompting the need to achieve agreement. Mechanisms and the ability to achieve alignment are therefore likely to become particularly important, especially for cultures that prize consensus. Practitioner Group considerations are explored further in a separate section.

8.1.2 Implications for MSS Problem Space constructs

The implications of the above argument for their existing MSS model mainly relate to what is termed the Problem Space, and repositioning the MSS in relation to the broader logics being enacted. These are addressed in turn below, first considering existing constructs that can be refined before considering missing constructs that may be important to include. To provide a point of reference for this discussion, an excerpt of the MSS model is depicted in Figure 8-4, showing the Problem Space causal constructs and related response constructs. See also Table 2-1 for descriptions.

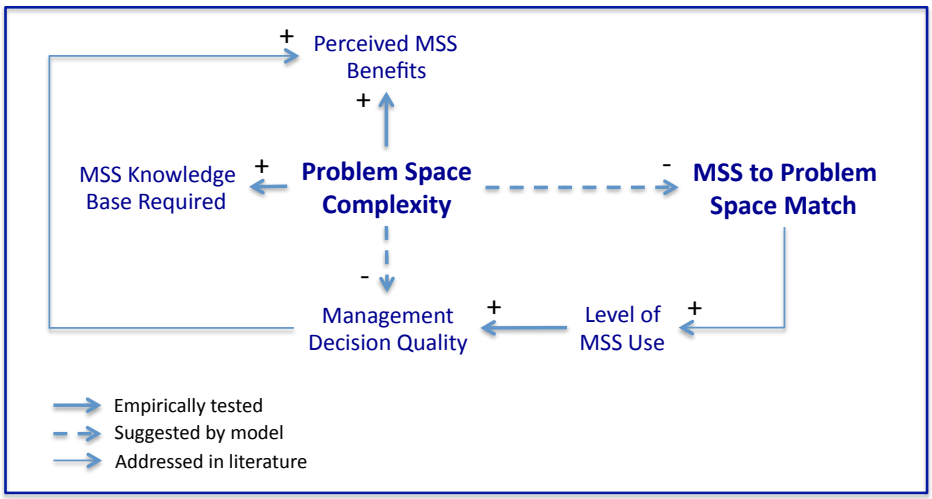


Figure 8-4 MSS Excerpt showing Problem Space related constructs
(based on Clark et al.: 2007)

Problem Space constructs are underdeveloped

The initial MSS critique in section 2.2.3 highlights that while several constructs related to the MSS Problem Space are described, a clear definition of Problem Space is not provided by the authors. This points to the need to more clearly define this concept, which essentially represents the context for inquiry and related use of any MSS.

The emerging explanatory framework presented in Chapter 7, and the related discussion of underpinning logics outlined above, both essentially address the practitioner context for pursuing insight and related benefits. Indeed, existing MSS ideas of the Problem Space and related constructs come closest to addressing the practitioner inquiry element at the heart of Data initiatives. These are acknowledged as critically important for MSS success, and my research observations certainly support this for exploratory initiatives.

Clark et al. (2007) also call for frameworks to address Problem Space Match and highlight that addressing Problem Space Complexity is an important dimension for achieving Problem Space Match. They characterise Problem Space Complexity as reflecting dynamic changes *external* to an organisation and related to identifying a wide set of relevant variables. They cite changes in customer behaviour and innovation as examples of such variables. However,

my research observations suggest a richer view of the Problem Space construct, and the need to more clearly define or decompose this construct in order to better theorise such initiatives and identify relevant Complexity and Match dimensions. Clearer delineation between Problem Space and wider context will avoid potential blurring. Here, making inquiry and benefit realisation logics explicit may help do so by defining what is in scope for the initiative and what is external to the project.

The explanatory framework brings into focus key constituent elements of the Problem Space more clearly: the Phenomenon that is the focus of inquiry, as well as the practitioner context, Knowledge and related benefits in view. The earlier discussion also highlights that Problem Space Complexity needs to take account of different sources of such Complexity. These could be reflected as different dimensions of Complexity, e.g. factors related to the Phenomenon, such as the current state of current practitioner Knowledge about it, the scale of aspiration that the desired new state of Knowledge represents. These could also be differentiated from the Complexity of the practitioner context, e.g. the scale and complexity of dependencies involved in realising inquiry or benefit objectives, powerful interests inclined to maintaining the status quo, etc.

These ideas are summarised and illustrated in Figure 8-5. Instead of an implicit Problem Space, distinct but overlapping Inquiry and Benefit Spaces are introduced, to highlight their interdependence. The demarcation, degree of overlap and of interdependence between Benefit and Inquiry Spaces depicted may also vary for different initiatives (as indicated by the arrow). For example, in some cases practitioners' activities may combine direct engagement with the Phenomenon, related incidental learning and making related practice improvements. The variation in these Spaces are indicated by using dotted lines as borders, and their overlap and interdependence by shading an indicative area of overlap. The different sources of Complexity within these spaces are reflected as separate Complexity 1 and Complexity 2 constructs for those associated with each. These Spaces are also differentiated from the wider context for the initiative or project.

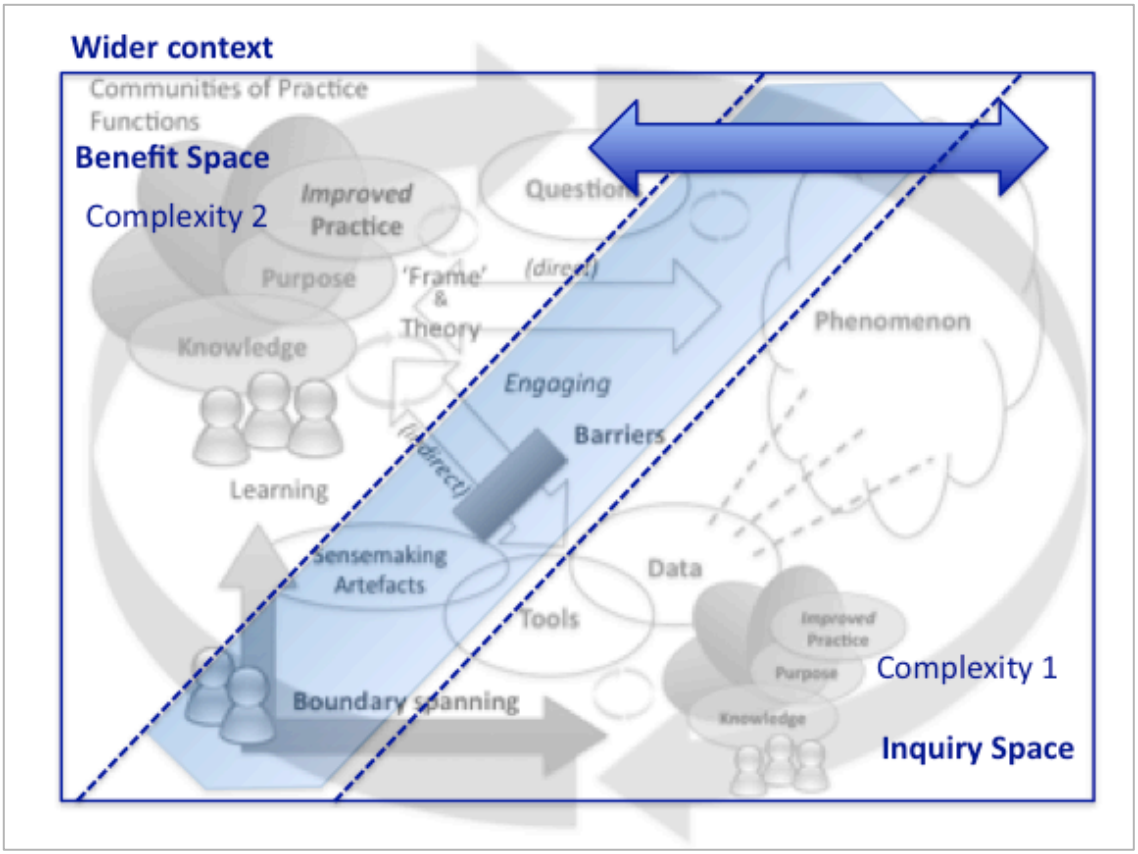


Figure 8-5 Benefits logic elements that frame the inquiry logic

More investigation is required to define a clearer Problem Space and related Problem Space Complexity constructs than those currently offered by the MSS model. These should aim to be easier both to measure and compare different initiatives, and to be more useful to practitioners in terms of suggesting appropriate interventions to better address different dimensions and states of Complexity. For instance, recognising distinct inquiry and benefit logics at work, points to the potential usefulness of separate related Complexity constructs and measures for each. This may be particularly useful where contexts for inquiry and benefit realisation may differ, e.g. where different Practitioner Groups are involved in inquiry elements versus application of insights to generate benefits, as in the InfraDig case.

Addressing Problem Ambiguity and Uncertainty

A further dimension of Complexity for Data initiatives relates to the level of clarity and certainty about the nature of the Problem or inquiry and the clarity

and confidence in the approach or logic being proposed to address it. Schön (1991/1983) argues that in practice "problems do not present themselves to the practitioner as givens" (p.40) and that we are inclined to ignore problem *setting* and the associated practitioner involved. Weick (1995) in turn refers to this in his broader framing argument for the role and importance of sensemaking in relation to such activity.

Various challenges in framing Problems

Several ideas noted in the literature review in Chapter 2, point to potential framing and related sensemaking challenges. For instance, Weick (1995) introduces the idea of focal narrowing, i.e. concentrating on the urgent and immediate under extreme pressure, which reduces sensemaking capacity. This loss of peripheral sensitivity may be important in conditions of (less extreme) pressure too. Paying attention to a wide set of potentially relevant cues, as well as wide consultation and involvement may be important to make more extensive connections to prior Knowledge and bring them to bear on ambiguous situations (Tsoukas: 2009, 2005).

This connects to ideas outlined earlier – of path dependency (Cohen & Levinthal: 1990), focal and peripheral attention (Polanyi: 1966) and bringing tacit Knowledge to bear subconsciously. Kahneman's (2011) cognitive notions of fast, instinctive thinking, versus slow, more considered or reflective thinking are also potentially useful. The former seem particularly likely to be impacted on by inherent cognitive biases. Tsoukas (2009, 2005) draws attention to the sensitising role of a person's 'inherited background Knowledge' to frame and pay attention to potentially relevant cues. This 'inherited background' is also constantly enriched by ongoing experience.

Distinguishing ambiguity from uncertainty

Weick (1995) makes a particularly important distinction between situations that are ambiguous rather than uncertain, arguing that sensemaking is required for the former, while more Data are likely to help the latter and may actually heighten confusion in the former situation. While his distinction is outlined in Appendix E.2 in more detail for reference, it is important here as exploratory

inquiry is more likely to require more Problem-setting effort addressing ambiguity, rather than simply addressing uncertainty. In this respect exploratory data initiatives may also be more Complex than those addressing uncertainty by providing more Data. Some Data projects may well include elements of both or lie on a continuum of such Complexity.

Theory Quality or Maturity

These issues are likely to crystallise when surfacing and articulating inquiry and benefit realisation logics. Therefeore the level of specificity of both inquiry and benefit realisation logics articulated and the level of confidence in them, e.g. in terms of clarity, prior use or experience, may represent a further dimension of Problem Space Complexity. However, it may be more useful to identify such Complexity as a separate consideration or factor entirely, perhaps as Theory Quality or Theory Maturity. I've avoided labelling this as Problem Clarity for reasons that will become clear in the next section.

At one level, Theory Quality or Maturity can relate to specific elements of either the inquiry or benefit realisation logic in terms of their associated level of specificity, or to the level of uncertainty or ambiguity associated with them. This is closely related to assumptions in the logics being used, and the sensitivity of the rest of the logic to changes in these assumptions, particularly of envisaged benefits. Addressing this may require exploration, validation and testing, as well as prioritising such clarification and testing effort before progress can be made. To the extent Theory Quality or Maturity complexity relates to the inquiry logic, addressing this may require more iterations of engagement with either or both the Phenomenon and data to clarify or test them further. Given the likely impact on resource requirements and time, where more effort and iterations are required, a Theory Quality or Maturity measure could be useful for evaluating investment proposals and project plans for exploratory data initiatives. This construct is illustrated in Figure 8-6, with separate elements positioned in relation to Benefit and Inquiry Spaces, reflecting the different logics involved.

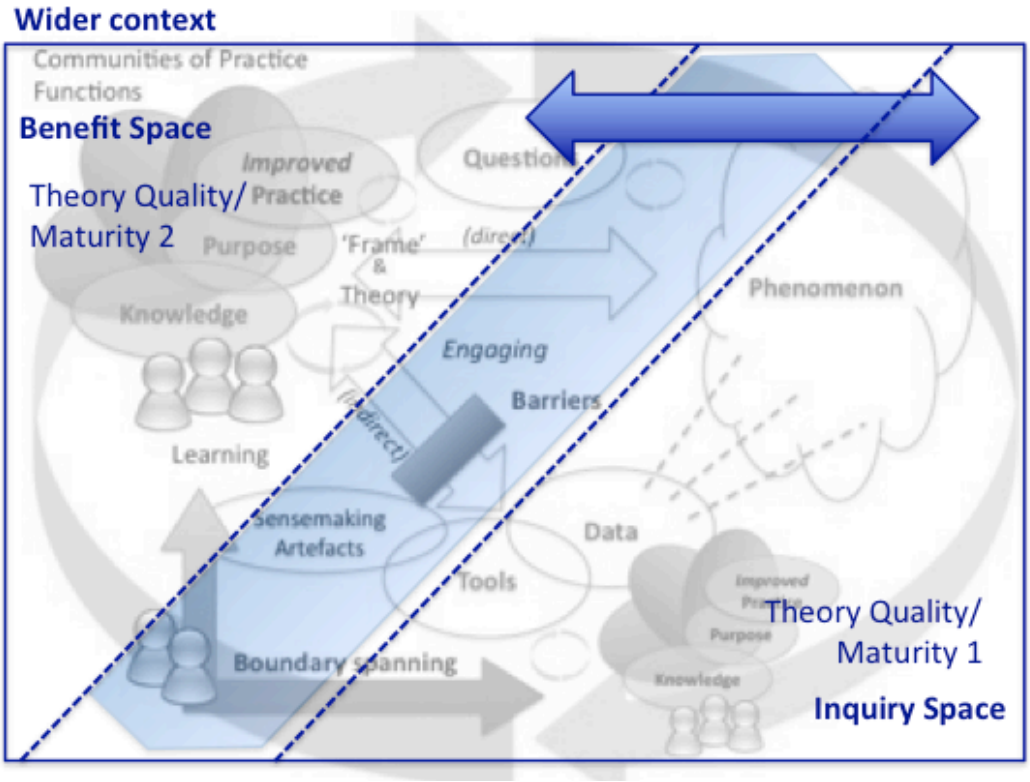


Figure 8-6 Benefit and Inquiry logic Theory Quality and Maturity

Ambiguity may also arise in relation to which options to pursue in terms of lines of inquiry or in terms of actions to take based on insights gained. This may prompt the requirement for sensemaking (Weick: 1995) and point to the greater likelihood of governance challenges to achieve alignment and consensus on courses of action, or indeed on desired benefits in view. Such alignment is revisited later when considering different Practitioner Groups and related institutionalised thinking. Questions of ambiguity about a Phenomenon may also point towards the initial value of direct Engagement before progressing to indirect Data Engagement, especially if ambiguity is likely to arise from how to interpret related Data. This seems closely associated with the idea of undertaking exploratory pilot research before embarking on more extensive Data collection and Data analysis in research settings.

A Problem Space framing is itself problematic

At a broader or second order level, a 'problem' framing is itself problematic, potentially introducing a framing bias for participants, managers and researchers. This is likely to be heightened in ambiguous situations and

exploratory Data projects. Such a framing may not place enough emphasis or focus on creative opportunities for innovation; this concern is consistent with wider work on *strategy blindness*, where projects fail to realise strategic benefits envisaged in spite of successful systems implementation (Arvidsson et al.: 2014). Institutionalised practices and *cognitive entrenchment* of participants or key actors are identified as important factors giving rise to such strategy blindness.

Framing an initiative as 'problem-solving' may lead to a problemistic search for potential solutions. Drawing on the behavioural model of the firm, Davis and Olson (1984) highlight that a problemistic search tends to be fairly simplistic, starting with initial symptoms or the current solution, only expanding the search if the original search fails to find adequate solutions, giving rise to risks of satisficing. Elsewhere they also explain that problem formulation typically involves clarification and simplification, often employing the following strategies to do so: bounding the problem, focusing on changes or variance analysis, decomposing the problem into more manageable subproblems and focusing on controllable elements.

This can perhaps be characterised as a convergent or focusing orientation rather than a divergent orientation focused on generating a wide variety of creative options (Leonard & Swap: 2006); it may lead to a bounded context and limited set of requirements being identified. This may limit avenues of inquiry and solution search compared to a creative, innovation focused or divergent framing, or framing which consciously addresses both. Such bounding and limited options may in turn lead to underestimating the lead time and number of refining iterations likely to be required for exploratory data initiatives, or simply ignore promising lines of inquiry and related benefits. The use of more straightforward requirement specification and project approaches early on in exploratory initiatives may also reinforce this, as requirements and likely benefits may not yet be fully appreciated. Indeed, Marchand and Peppard (2013) argue for adopting more agile, iterative approaches for such initiatives, which would also help address ambiguity challenges. Therefore,

relabelling or reframing the Problem Space construct may clearly be useful to avoid such a bias and better reflect and encompass both convergent and divergent approaches to identify potential inquiry and benefit options.

This also brings the broader innovation literature into view, as potentially useful for future research and IS theorising in this area. For instance, it may shed light on the likely tension within such initiatives to balance a need for divergent and convergent thinking, to generate creative options on the one hand, versus exploring and prioritising which options to pursue (Leonard and Swap, 2006). Eisner (2003) also alerts us to the potential impact of organisational performance on framing and related solution search strategies that may be adopted. Where performance is assessed favourably, *opportunistic* search strategies are pursued, while *problemistic* search strategies are pursued in response to an adverse performance assessment. Attention is also drawn to the influence of prior positive performance entrenching strategies, though significant environmental change makes the performance link more ambiguous. This argument echoes the findings that cognitive entrenchment bounds initiatives, leading to strategy blindness (Arvidsson et al.: 2014). Similar considerations may well be relevant for exploratory Data initiatives that engage in search strategies and activities, especially if they are aimed at or support strategic change, or represent a response to external changes.

As already identified in the literature review chapter, the OL and KM literatures may provide further useful ideas for theorising in this area (Eisner: 2003, Gherardi: 1999,). However, many of these ideas are based on research focused at an organisational level, so related constructs and models may require careful adaptation and validation to evaluate if they apply in the same way at the level of initiatives or projects.

For example, Salge and Vera (2013) argue for the importance of slack resources for incremental learning, which they conceptualise as a dynamic organisational capability. On the face of it this may also be important for individual data initiatives, or Data orientation and Engagement more widely, if we characterise these as learning initiatives and activity. However, the concept

of slack resources may need to be recontextualised within the project or Practitioner Group environment, perhaps to encompass support and free time made available for boundary spanning or exploratory Data Engagement, as Tanya undertakes at GoCouncil in relation to Acorn Data.

Might Data project categorisation help?

Recognising problem, learning or creative imperatives for a project or initiative may therefore be important for framing and planning Data initiatives, and are likely to be reflected in their inquiry and benefit realisation logics. However, a neat categorisation may be challenging. Projects may well occur on a continuum, pragmatically including several elements that fall into either category, or lie at different points on such a continuum. Projects may also have different imperatives at different stages. Different inquiry and benefit realisation logic elements or steps may benefit from or indeed necessitate different problem or creative frames and imperatives. For example, considering the scope of Data gathering, or different options for Data analysis, may benefit from thinking broadly, versus actual Data gathering and executing analysis. Instead, a related, broad categorisation of Data initiatives in terms of their exploratory or exploitative Knowledge or theory focus (i.e. theory-building/refinement or application of theory focus) may well prove more useful. These categories of projects are likely to be characterised by very different inquiry and benefit realisation logics and throw up very different challenges, requiring different approaches which may tie into wider research on ambidexterity (Huang et al.: 2014, Turner, Swart & Maylor: 2013).

The above argument extends the closely associated argument for the need to reframe outcomes as broader than just decision-making benefits. This was already highlighted in the literature review (see section 2.2.4) and is borne out by my case observations.

8.1.3 Reframing Tool/MSS Use

The central focus on MSS implementation and the MSS' underpinning processing logic has already been problematised in the literature review.

However, the identification of separate but interlocking inquiry and benefit realisation logics at work in Data projects also provides a useful basis for reconsidering and recontextualising tool or MSS Use. It prompts us to consider how an MSS enables both the inquiry effort and the benefit realisation effort. As argued above, these seem fundamentally different requirements involving very different logics, challenging the notion of a single MSS Problem Space Match construct. In particular, the existence of two logics forces us to reconsider the concept of MSS Problem Space Match.

Figure 8-7 The MSS Artefact in relation to Framework Data & Tool elements below positions the MSS artefact within my explanatory framework, in relation to covering its Tool and Data elements (and related Sensemaking Artefact outputs). Doing so also positions the MSS in relation to Inquiry and Benefit Spaces, prompting the recognition of different Inquiry Space and Benefit Space MSS Match Constructs being likely, which is explored next.

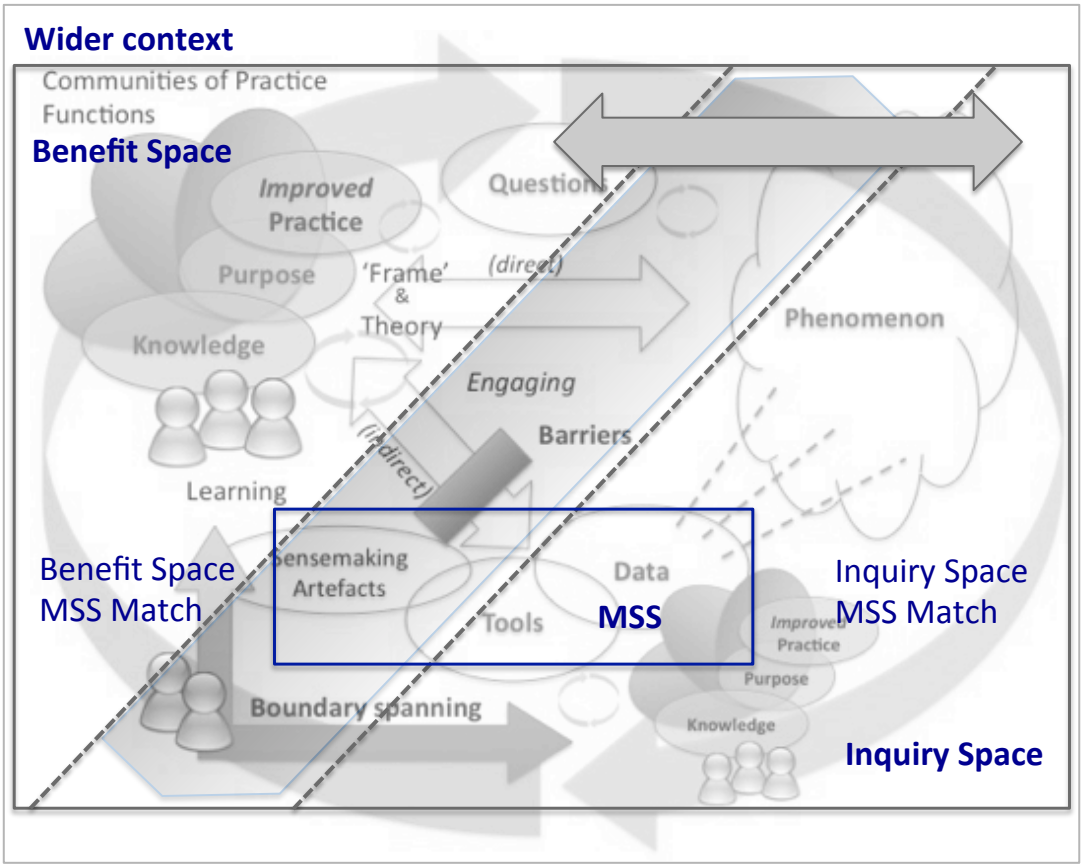


Figure 8-7 The MSS Artefact in relation to Framework Data & Tool elements

Positioning the MSS highlights different MSS Match requirements

As illustrated above, repositioning the MSS to correspond to Framework Tool and Data elements illustrates how an MSS may span and address both Inquiry and Benefit Spaces.

However, the very different requirements related to inquiry and benefit realisation also prompt a re-evaluation of the implicit assumption of a single or unitary MSS construct or artefact. As highlighted by the case comparison, a variety of Tools may be used rather than a single, all-encompassing MSS. The choice of Tool or Tools may reflect the tasks at hand and facilitate different kinds of Data engagement, e.g. Data collection may require particular Tools, such as barcode scanners or CAD systems in the case of InfraDig, or audio recording equipment in the case of research interviews. By contrast, Data analysis was undertaken with a variety of Tools in the case of my research. This points to a further possible distinction, between Tools related to Data collection (including sensor measurement and recording) on the one hand, and those related to Data organisation and analysis on the other. GoCouncil also illustrates that where secondary Data are acquired, it may make Data collection tools seem less relevant to the inquiry logic being pursued.

Based on the case experiences, Tools used to support Data collection can be relatively simple, in contrast to those for Data organisation and Engagement, such as eB application and database software at InfraDig, or NVivo and Prezi data analysis Tools in my research. Analytical Tools seem far more closely or directly associated with the inquiry logic, often encoding particular forms of analysis in algorithmic terms (e.g. SAS and SPSS statistical analysis tools). Such Tools would require considerable flexibility to support pursuing various strategies for Data exploration, experimentation and refinement of both Data and related analysis. More typically such Tools evolve to support more specialist analytical or research approaches, e.g. different quantitative versus qualitative Tools exist and are more often used for particular forms of analysis (Lewis & Silver: 2007). More generalist Tools, such as Excel or SQL by contrast have less sophisticated analytical functionality but may allow more flexible Data

engagement, implying a potential trade-off between depth and breadth of Tool functionality.

The scale of Data collection and analysis requirements may be a potentially important factor in determining MSS Problem Space Match in relation to Inquiry requirements. It also points to the possibility, indeed likelihood, of a suite of Tools being adopted to address a range of Data analysis strategies. For example, quantitative Tools might be required for the statistical analysis of quantitative Data, while other Tools are likely to be more appropriate to code and analyse qualitative Data (e.g. texts, video, audio). Therefore, Tool or MSS choices should reflect the scale and Complexity of the Data collection effort, the nature of the Data, and analysis being undertaken, as well as user experience and preference. These choices are further influenced by existing organisational Tool or systems investment and ready availability. A provisional, bricolage approach may result, using familiar and readily available end-user Tools, rather than a planned or designed approach using a single MSS. This may be particularly likely in contexts of individual analysts, where little coordination and standardisation may be required, or where Data volumes are modest. Such ideas are consistent with other KM findings in relation to practitioners adopting bricolage approaches (Duymedjian & Ruling: 2010, Gherardi: 1999).

By contrast, benefit realisation logics are likely to involve very different IT enablement or support requirements. They are more likely to be exploitative than exploratory in nature and may need to be integrated into existing operational systems. In the case of InfraDig, this is highlighted by the Operator requirement for Data to be readily available to field maintenance engineers via their existing operational maintenance and asset management systems. This may well require Data migration from the InfraDig eB asset Data system. Therefore, consistency with existing Tools and practices may also be more important to aid adoption or MSS Use and benefit realisation than for Data exploration. This reflects the likely scale of existing systems or Tool investment, and institutionalised practice, especially where far greater numbers of users or practitioners are involved.

An extract of the MSS model related to the MSS Problem Space Match and Use constructs is shown in Figure 8-8 as a point of reference. This highlights that MSS to Problem Space Match is cited as the main factor in MSS Use. The authors include an MSS' ability to deal with selecting and filtering a potentially overwhelming volume of Data for what is relevant to their Problem Space as an important aspect of such MSS matching.

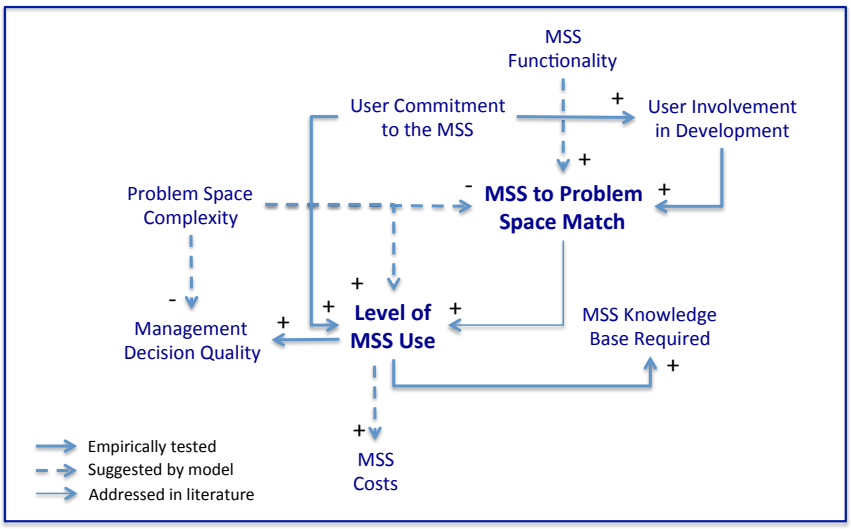


Figure 8-8 MSS Excerpt showing MSS to Problem Space Match and MSS Use related constructs (based on Clark et al.: 2007)

In order to promote better MSS Problem Space Match, several methods are acknowledged by the MSS authors, which all have an iterative quality to better tease out envisaged use and related Data requirements. In their related discussion, the MSS authors also recognise that Data collection, organisation and Knowledge foundations can represent both a 'bottleneck' and impair outcomes. They also recognise a call to integrate DSS and KMS to help address matching and also argue for the importance of MSS Training and related MSS Knowledge. This connects to path dependency and absorptive capacity ideas highlighted earlier (Cohen & Levinthal: 1990).

The case observations certainly provide support for the importance of prior Knowledge and the investment hurdle represented by Data collection, as a potential Barrier to Engagement. Clark et al. (2007) acknowledge a call to integrate DSS and KMS, which points to the importance of wider KM and

learning. However, this tends to imply and emphasise existing codified Knowledge rather than relevant tacit Knowledge and the more emergent learning observed and experienced. This is revisited when discussing human elements that should be recognised within the model.

Data Engagement rather than MSS Use as central to insight

The earlier discussion of MSS Problem Space Match reflects how Tool or MSS Use or adoption may reflect a variety of factors rather than simply MSS User or Management Commitment to an MSS, and MSS or Tool related Knowledge, as currently identified in the MSS model. More fundamentally, the case observations and research experience highlight how Data Engagement is central to participants pursuing and generating insight from Data, however this is achieved and whichever combination of Tools is used. Therefore, a better concept of MSS inquiry related Match for any particular MSS or combination of tools may be how well it enables Data Engagement.

However, in the light of this, a more radical MSS model revision may be warranted, to adopt the quality and extent of Data Engagement *instead* of MSS Use as a core construct and factor within an MSS model. Such a construct might reflect the number of iterations and variety of Engagement strategies or angles adopted as further indicators or measures of the quality of Data Engagement achieved.

A primary emphasis on Data Engagement rather than MSS Use will also allow the resulting model to hold for instances where no MSS or Tool is used. The model would then better address the wider MSS Information System, rather than the narrow MSS IT artefact and its use. Such a focus on the wider IS is argued for by Lee, Thomas and Baskerville (2015) in the context of systems design science, following in a long tradition of earlier sociotechnical and systemic thinkers (Checkland & Holwell: 1998). Indeed, this highlights a relative paucity of IS research about Data use and users within IS, even within the data analytics context (Tamm et al.: 2013). In particular, there seems to be little broad research on patterns of Data use, the nature of related Information users generate from Data, or the sense they make of it, and how they do this in

different practitioner and organisational domains and contexts. This may reflect our traditional failure, as a discipline, to really distinguish and engage with the inherent nature of Information and Knowledge as ineffable and human, as highlighted earlier in section 2.8.1. It may also lie at the heart of the continuing challenge to consistently realise expected IS outcomes and benefits from related investment.

Figure 8-9 seeks to address and illustrate the above by giving central prominence to the Data framework element encompassed within the repositioned MSS construct. It also highlights the related Data Engagement element as central, recognising new constructs related to the Quality of such Engagement, as well as potential Barriers to such Engagement. These are explored next.

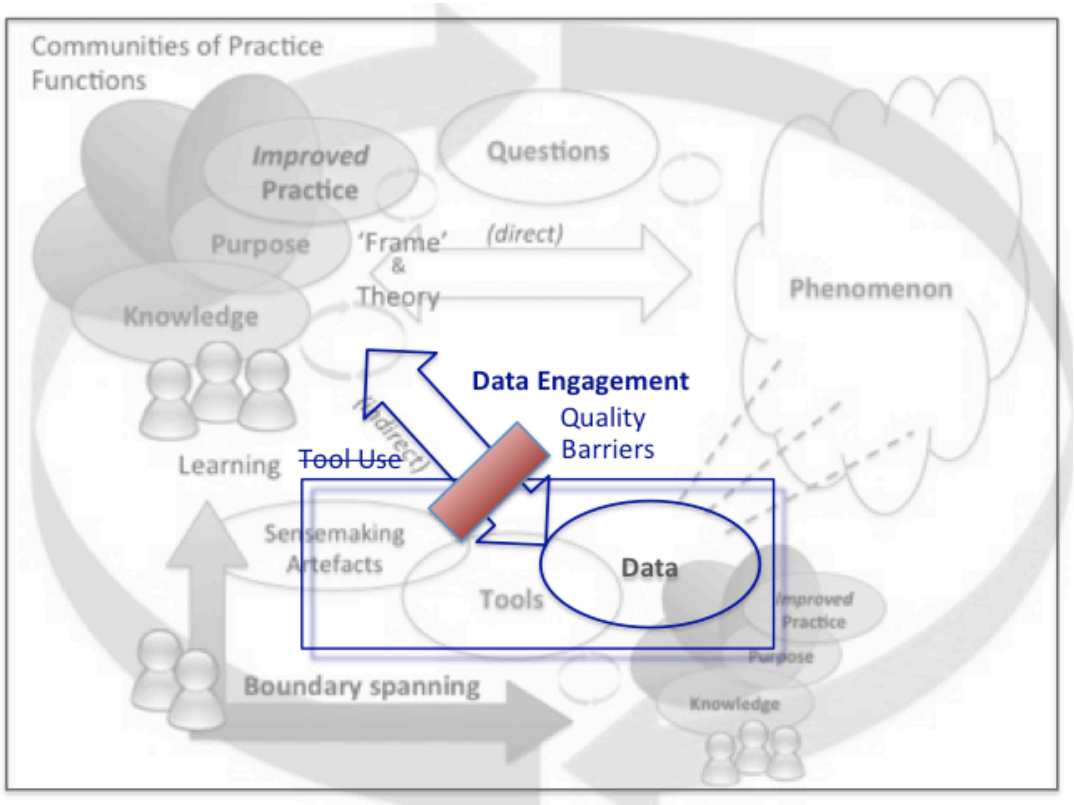


Figure 8-9 Emphasing Data Engagement as central rather than MSS or Tool Use

As highlighted above, my research observations point to the existence of a number of potential Barriers to Data Engagement that may be important (see section 7.7). Some Barriers are closely related to the MSS itself. For instance,

given its role in mediating Engagement with Data, an MSS or Tool may well bound or limit Data Engagement in a number of ways: what Data can be captured, what Data are accessible, how they can be represented, what analysis is possible, etc. Other Barriers may relate to institutionalised practice, social and organisational factors, for example Data orientation (Marchand et al: 2001). Therefore, the idea of Barriers to Data Engagement represents an important additional construct. Such an approach might also reframe MSS Knowledge as a potential Barrier to Data Engagement rather than an essential enabler of MSS Use. Further research across different Data project settings may well identify further Barriers, and indeed identify enablers as well.

8.1.4 Give Data prominence and disentangle it from wider MSS

Extending the reframing of MSS artefacts or Tools used outlined above and in section 7.6, this section examines the MSS artefacts and tools used themselves. It does so with a view to better theorising the MSS artefacts used within the context of a Data project or initiative, with a particular focus on Data.

Given the centrality of Data Engagement to produce insight already highlighted, it is striking that the MSS model does not include an explicit Data construct. Instead it is simply considered part of a wider MSS artefact. Nevertheless, the importance of an MSS' underlying Data is touched on as important in a number of areas within the MSS model. For example, the discussion of the MSS Training construct acknowledges a need for training related to its underlying Data model. It is also indirectly reflected in the discussion of MSS Usability, where the models used, in particular by DSS systems, are acknowledged as an important and often overlooked aspect of MSS Quality.

To address this MSS model shortcoming, this section seeks to build on the following existing efforts within IS and related arguments outlined in my literature review (section 2.8):

- Efforts to conceptualise Data, Information and Knowledge, and
- Sociomaterial efforts to theorise the IT artefact

The discussion starts by introducing a practical generic example of an IT system as a reference point for the discussion. Based on my case observations it then goes on to:

- Present an argument for disentangling Data, distinguishing it from algorithmic elements or components, when theorising IT artefacts
- Outline several characteristics or features of Data that emerge from my research as important considerations in defining and theorising such a Data IT artefact

The section concludes by returning to some preliminary implications for the MSS model. As well as a distinct Data artefact construct, some related constructs are introduced that may be useful. It is recognised that these are necessarily provisional and that further research is warranted in this area.

Introducing a generic system example as a reference point

Withing IS practice and typical systems design, Data are often seen as distinct and a separate constituent element, e.g. databases. I've sought to illustrate the common basic software and hardware architectural elements (including sensors) encountered in practice in Figure 8-10, also reflecting the users and designers involved in related activities. Figure 8-10 illustrates how systems use algorithms to capture, process, store and produce Data for use. Programs or algorithmic elements are typically held separately in programs and 'applied to' or 'interrogate' related Data, typically stored in separate databases or datafiles.

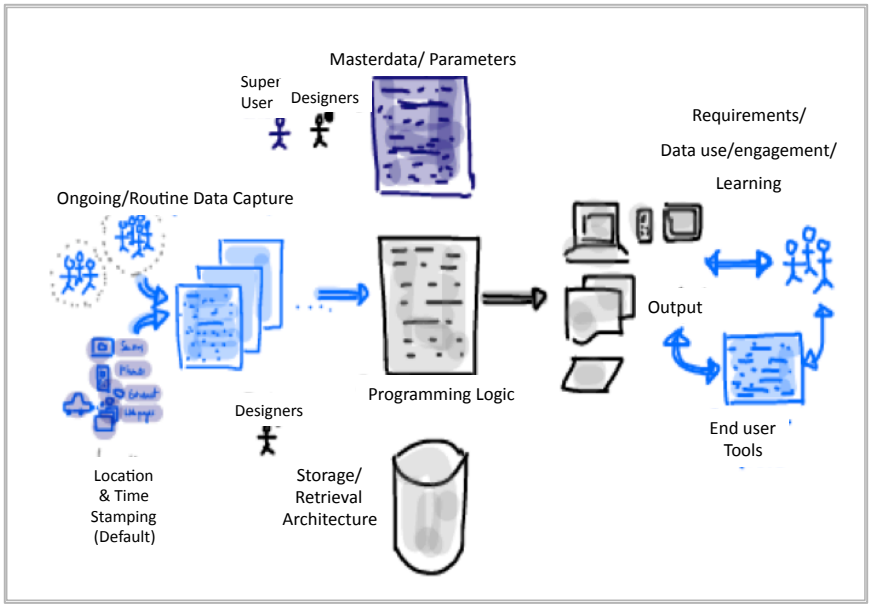


Figure 8-10 Common basic software and data architecture

While considerable variations on this basic architecture exist, reflecting different distributions of these elements, recursive layering of software (e.g. database, operating system, communication layers), the broad distinction between Data elements and programmes persists. Indeed big Data practitioners often stress the separate or independent nature of Data, distinct from related systems and Tools. At the same time, the above reflects the basic and long held IT systems' conceptual model for Data processing: the input, process/storage and output model (Davis & Olson: 1984).

The argument for including distinct Data and Algorithmic IT artefacts

The argument for recognising a separate Data construct or artefact within the MSS model, distinct from algorithmic and wider elements, has several facets. Some facets are pragmatic and aimed at improving the MSS model's normative value, while others are aimed at better conceptualising and theorising IT artefacts. Such a distinction may indeed be useful when theorising more general IT artefacts in other contexts also.

The following facets are briefly explored in turn:

- Ensuring Data characteristics and dimensions are not overlooked
- Data and algorithms are likely to encode different Knowledge
- Data and algorithms can evolve differently over time

Ensuring Data characteristics and dimensions are not overlooked

InfraDig highlights how Data are often embedded within or tightly coupled to IT artefacts, particularly application software. It also illustrates that the degree of such Data embedding varies for different tools or applications. An emerging trend towards 'Data-fication' of most technology artefacts may even blur the distinction between IT artefacts and broader technology artefacts.

The lack of separation between Data and algorithmic elements of the MSS IT artefact overlooks Data's unique characteristics, when compared to algorithmic elements. This results in important related activities and factors, such as Data Engagement, remaining hidden from view and therefore under-theorised. Reviewing Figure 8-10 highlights other missing areas, including Data capture or acquisition, and an associated Data Quality construct (distinct from MSS Quality), Data organisation and related design reflected in database structures, technical formats, taxonomies, categories, etc. Some of these aspects emerge as important in the InfraDig case and also in my own research coding experience. Practitioner groups (ISACA: 2013, DAMA: 2009) highlight further aspects too and these are by no means meant to represent a comprehensive set of relevant Data related considerations.

These concerns seem distinct from Data processing, which applies algorithms to Data for analytical, transformative or communicative purposes. As illustrated in Figure 8-10, representational considerations, e.g. formats, are worth distinguishing from 'content' related questions more aligned to user meaning. This is reflected in some of the conceptual groupings of attributes identified in relation to Data and information. Some practitioners (ISACA: 2013) identify broad intrinsic versus contextual (and non-functional security and availability) groupings of attributes. Some researchers (Boell & Cecez-Kecmanovic: 2010)

go further, arguing for distinguishing up to seven different conceptual layers and related information attributes, from physical through to social, based on an extended semiological framework attributed to Stamper (1991). However, these approaches still don't address or offer a clear basis for distinguishing between Data and information as concepts or constructs.

Separating Data and algorithmic elements may also help us understand and theorise MSS evolution more clearly. For instance, my research coding experience illustrated how my Data sometimes evolved independently of the Tools used (e.g. coding elaboration). While Tool or Data evolution is not explicitly reflected in the current model, the MSS authors do acknowledge that development of MSS and related user involvement is often iterative, and that iterative learning about an MSS occurs through use. Shifting requirements and use of MSS are also noted in their discussion of the MSS Functionality construct. For them, this highlights a need for MSS flexibility, to address different problem stages, as well as the need to accommodate changes or to rapidly develop the MSS in response to changes in the Problem Space.

However, my research highlights that the whole endeavour is fundamentally iterative in nature and at various levels. In addition to evolving at the data element or field level, as described in relation to coding elaboration, this may also be reflected in a changing set of Tools used rather than changing the use of a single MSS. This may in turn require less flexibility from a particular MSS or tool, and rather a flexible and changing set of Tools or range of (compatible) Tools available. However, most importantly using different Tools in relation to the same Data also underlines its independent and distinct nature.

Data and algorithms are likely to code different Knowledge

My research observation and emerging explanatory model point to different kinds of Knowledge being incorporated within Data versus within algorithmic elements. This is explored below, focusing on user contextual meaning and functional purpose rather than representational aspects, such as format or syntax.

As explained in section 8.1.1, Data always exist in relation to a particular Phenomenon or Phenomena. Data elements, typically fields within a dataset, reflect dimensions about Phenomena deemed relevant during system design, while field entries may reflect categorisations or distinctions identified at the time, or state measurements. Tsoukas (2005) argues that the ability to draw increasingly fine distinctions about a Phenomenon is a good basis for defining knowing in relation to a particular Phenomenon. In this sense datafields and categorisations reflect Knowledge about the Phenomena the Data seek to represent. In the case of relational database structures, this also reflects system designers' understanding about relationships between Phenomena or Data entities, related dimensions or datafields, and content states or classifications. However, such structures may also reflect their assumptions and Knowledge in relation to how best to organise Data for storage, processing or presentation for intended or envisaged use, rather than just in terms of Knowledge about the Phenomena in view. Design choices about the use of relational, hierarchical or 'flatfile' Data structures are also more likely to reflect processing and storage considerations than Knowledge about the Phenomena being represented.

By contrast, algorithmic elements seem primarily concerned with the efficient, optimal storage and processing of Data. To the extent they encode Knowledge related to particular Data analysis techniques, e.g. statistical routines, this still seems like a specific case of Data processing Knowledge. However, this picture shifts when considering algorithms related to benefit realisation. In such instances they may well encode business rules, predictive models and other Knowledge or theory related to particular Phenomena, for example, rules about when to undertake particular maintenance activities or part replacements.

Furthermore, where algorithmic elements address the resulting presentation of processed Data for users, e.g. in the form of reports or user display, this may reflect Knowledge about the context of practitioner use. For example, it could reflect their Data Engagement preferences, broader Purposes and activities, physical location and conditions (such as access to light, power and network availability) Clearly these can relate to both inquiry and benefit realisation

requirements. As seen in the InfraDig asset Data team in terms of eB, users often adopt familiar end-user systems and tools to address systems' shortcomings in this area, exporting Data where possible, to be able to tailor their Data Engagement and presentation. While the recent recognition of the importance of visualisation (Chen et al.: 2012) for Data exploration may prompt software vendors to pay this aspect more attention, its design seems to warrant greater attention.

Data and algorithms can evolve differently over time

Clarifying the different kinds of Knowledge that may be reified and optimised in an IT artefact facilitates focusing on how these elements and related Knowledge interact and (co)-evolve over time. Importantly, Knowledge about Phenomena may evolve at a different rate, and in response to very different triggers, compared to Knowledge about how best to organise, process, store and present Data. Knowledge about how practitioners may best use and apply Data collected in a practitioner context may also be worth distinguishing.

For example, Data may evolve in response to changes in the Phenomena of interest to practitioners, prompting the need for new datasets or datafields; their growing understanding of particular Phenomena, e.g. in terms of more elaborate field categorisation, reflecting finer distinctions being drawn (Tsoukas: 2005); or previously unrecognised relationships, resulting in new entity or field relationships emerging, requiring Data reorganisation. This became personally apparent through my own Data analysis, when after my coding stabilised a need emerged to reorder my Data along a timeline. It was also somewhat evident within InfraDig during the discussion of KPI Data requirements, though more so in the discussion with OtherCouncil, who indicated that their index gradually stabilised over a two-year period as dimensions and measurement of deprivation in different contexts of use became clearer. These examples also highlight how Knowledge and related Data may evolve more quickly for new or unfamiliar Phenomena than for those that are well-understood. By extrapolation, this is also more likely to be true for rapidly changing practitioner contexts than for stable practitioner contexts.

By contrast, Data processing logic is likely to evolve differently, depending on its inquiry or benefit realisation focus. For inquiry related processing, this is likely to be prompted by new analytical techniques emerging, more efficient processing algorithms or new presentation techniques. These may not be as likely to be context- or Phenomena-dependent and the scope of their impact and benefits of adopting them may also be more pervasive. Data processing supporting benefit realisation or execution logic, on the other hand is likely to change in response to changing Practitioner Group concerns and related practice improvements being sought. These may involve associated changes in business rules, reflecting changing priorities and external developments, as well as different processing optimisation requirements (e.g. timeliness).

Interrelated changes and opportunities might also be better anticipated, evaluated and accommodated with a clearer separation of algorithmic and Data elements, while keeping both in view. Doing so avoids design ‘lock-in’ in one area, hampering or bounding evolution in another. One can certainly foresee an external environmental change prompting a need for new Data as well as new business rules, for example the new business models being envisaged by GoCouncil. Similarly, new Data becoming available about an existing Phenomenon of interest may prompt changes to Data elements, Data relationships or structure. These represent progressively more complexity and impact in terms of making such changes. However, a new type or form of Data altogether also presents new processing and analytical challenges. For example, using visual or audio Data in contrast to textual Data presents very different Data processing and storage challenges.

Such differential evolution is highlighted visually in Figure 8-11 in relation to the different elements of the typical technical architecture encountered in practice, presented earlier. Variation is indicated by shading elements differently, also adding a related circular key indicator.

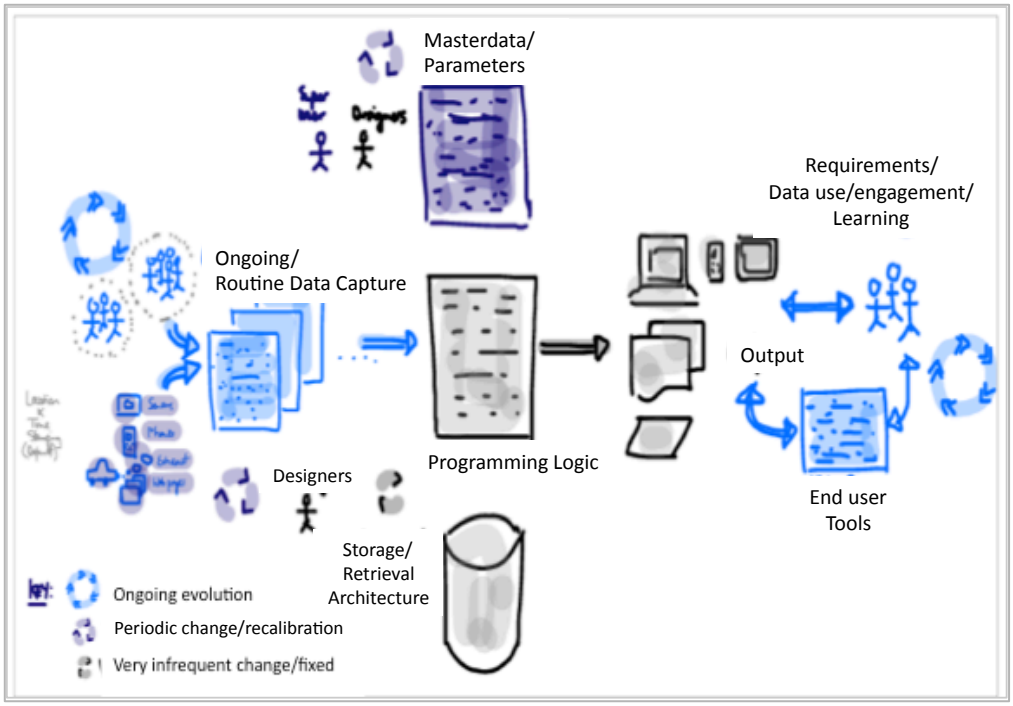


Figure 8-11 Differential evolution of IT artefact elements

Figure 8-11 shows how elements of an IT artefact differ in their evolution. Some examples of different triggers that may prompt such changes are also provided in relation to each element below:

- **Routine Data captured** – from sensors or via human observation and Data capture staff of related Data (e.g. forms, input screens). This is typically an ongoing activity of human sensemaking, measurement and categorisation. Where automated via sensors, changes are likely to be less frequent, requiring overall system design changes and technology changes, as for algorithmic elements below.
- **Master data and parameters** – configured by designers and/or senior managers. This is normally captured initially at systems design and implementation, and periodically thereafter.
- **Algorithmic programming logic and related digital storage and reporting.** These tend to change relatively infrequently and in concert. Such change is often dependent on the application provider, or systems designers and software developers, and may be prompted by changes in underlying technology used rather than evolving user requirements.

- **End-user-computing tools** often used by users to facilitate more flexible Data Engagement and further analysis, as well as tailoring Data presentation for particular purposes. Such Engagement, analysis and presentation are likely to evolve with changing practitioner requirements, reflecting evolving practitioner goals, challenges and related questions.

The figure also seeks to show the different authorship of particular elements, which emerges as an important consideration in theorising the Data artefact and is explored in the next section.

Particular considerations for theorising Data artefacts

The previous section argued for the value of disentangling Data as a separate element of a wider IT artefact, or indeed as a particular kind of IT artefact. This section explores several considerations that emerge from my research as important to address when conceptualisng and theorising such distinct Data IT artefacts, which are explored in turn:

- Keeping authorship in view
- The need for making distinctions between data processing and human sensemaking and keeping both clearly in view
- Differences and shifts in the related balance of human and material agency over time
- The value of a practitioner group view of authorship over time
- Accounting for the boundary role Data can play in some contexts

This discussion assumes a sociomaterial starting point for such theorising and some research observations also have implications for current debates about theorising IT artefacts within the sociolamateriality discourse, as outlined in section 2.8.2.

Keeping authorship in view

A central debate within sociomateriality relates to treating humans and IT artefacts as straightforward, equivalent actants. In treating them as equivalent actants, there is an attempt "…to allow the technology to speak for itself…" (Cecez-Kecmanovic et al.: 2014: p.826). Beyond the dangers of implicitly

slipping into anthropomorphising artefacts, the design authorship, and related objectives and context, or starting conditions, for such artefacts may be lost from view. This is recognised to some extent in an acknowledged and explicit emphasis and focus on performative aspects of a Phenomenon in such research, i.e. focusing on how rather than why things occur, when adopting such actor network approaches.

Given the considerable Knowledge that is inscribed or encoded by designers, losing sight of such authorship may significantly hamper our understanding, addressing why questions and related theorising. In particular, researchers adopting such approaches may not pay adequate attention to how practices and related practitioner Knowledge may evolve compared to more reified or rigid system encoded Knowledge. Important resulting tensions, constraints or accommodations enacted over time may therefore also be missed.

My research observations and discussion richly illustrate how economies of meaning, their ownership and negotiation are important aspects of how Practitioner Groups make their Knowledge and learning explicit, allowing this to evolve. It also highlights how they can represent Barriers to Data Engagement for other Groups. Such features should clearly be an important part of any theorising about IT artefacts or systems as reifications of such Knowledge and meanings. This is likely to be especially important when such IT artefacts represent Boundary Objects. Keeping the systems' authorship in view also allows inscription, dialectical and re-interpretive acts or moves to be examined over time. A narrative description or approach also ascribes any motives, constraints and related systems design decisions to the agents or authors that are their source, rather than treating them as an inherent feature or characteristic of an artefact or actant.

Recognising and distinguishing clear authorship in our theorising may also allow us to more clearly attribute and recognise Data and intellectual contributions encoded in systems. This could have important implications for how related benefits and rewards are shared, incentives are designed, or indeed prompt

greater scope and sophistication for protecting and transacting encoded intellectual capital.

Keeping clear authorship in view is also closely related to important questions of Data privacy, ownership and governance. For example, users who provide Data via devices, when they use applications or browse, may have reasonable expectations of what represents appropriate use, whatever the terms of use they implicitly are forced to accept when downloading related software. This is becoming an increasing focus of both privacy activists and regulators, and a clearer view of Data authorship could enable and inform related debates, policy decisions and managing such issues.

Distinctions between human sensemaking and Data processing

My research illustrates how Data can clearly exist separately as artefacts across space and time, e.g. Acorn Data that can be acquired and used by GoCouncil. On the one hand, this supports Leonardi's (2013) argument for making a distinction between human actants and artefacts, which persist independently over space and time, independent of humans. On the other hand, across both cases, my research illustrates how Data both reflect and inform practitioner contexts and co-evolve with Knowledge about Phenomena the practitioners are interested in.

Sensitivity to system elements that remain fixed, versus those that enable natural practical evolution and reframing, seems an important consideration for theorising IT artefacts, and for the practical design of such systems and related Data. It highlights the important and distinctive ability of human agents or actants to constantly reframe and evolve new meanings in response to changes in their practice, context or purpose. This represents a fundamental difference to most IT artefacts or actants. IT artefacts represent or incorporate codified, and by definition, bounded algorithmic and structural logic, focused on requirements in view at the design stage. Even learning systems currently operate within some framing constraints and assumptions.

While reinterpretation and repurposing of IT artefacts and underlying logic is clearly possible and likely during use (Orlikowski: 1992), my research illustrates how this may be bounded and constrained and will also depend on the flexibility inherent in the IT artefact. Its inherently codified and bounded nature is likely to result in many similar shortcomings and challenges to those highlighted in relation to codified Knowledge more generally (Tsoukas: 2005), especially when failing to recognise tacit precepts or assumptions. Weick (1995) raises concerns about codified systems not being able to adjust and reframe in the light of changes or encountering unanticipated situations. Star (2010), similarly alerts us to widespread category problems, when things don't quite fit the categories available.

This clearly points to the important consideration of the level and intensity to which a given IT artefact, or class of IT artefacts, both encodes Knowledge (and related Data) and imposes such Knowledge, either implicitly through its mediation (e.g. presentation, user interface or options), or explicitly enacts it through its Data processing rules and Data output. One could envisage that this might be a more important consideration for IT artefacts or systems that apply complex algorithms and processing, i.e. enacting related business rules and codified Knowledge, compared to those that facilitate communication or Data Engagement and exploration. In sociomaterial terms, this may well represent an important aspect of the agency attributable to an IT artefact in use. This seems an interesting line of inquiry and an important consideration when theorising IT artefacts.

The different evolution of various IT system elements outlined in the previous section also clearly has a significant impact on designing related MSS components, and how tightly these could or should be coupled. A re-evalution of MSS concepts of MSS Functionality and flexibility, as well as MSS Quality and MSS Complexity assessment is required to take account of this, ideally for each of the different components outlined.

My observations also highlight the likelihood of a trade-off between flexibility and optimisation in any IT artefact design. They point to an ongoing tension

between these objectives, arguing for the need to manage the tension and guard against optimising early on in Data initiatives, especially during exploratorion. Optimising may be more appropriate for IT artefacts used for exploitation, once Data and Knowledge about a Phenomenon have stabilised, and likely Data use is also clear. Even then it may require periodic re-evaluation and, where necessary, recalibration or even more fundamental re-architecting, to address evolving requirements and reflect ongoing learning during use.

A measure for the level of change or evolution in component Knowledge encoded or reflected may be an interesting area for further investigation. On the one hand, this may provide a rough indicator of the level of learning or innovation occurring in relation to the encoded Knowledge, which could in turn be benchmarked or compared across contexts or Knowledge domain areas. On the other hand, while a lack of change or slowing may reflect maturing Knowledge or environmental stability, it may also provide an early indicator of institutionalised thinking or design lock-in. Such indicators could in turn serve to identify potential areas to target or prioritise for improvement, experimentation and innovation.

Differences and shifts in the balance of human and material agency

Earlier I cautioned against the danger of masking or losing sight of the role and agency of designers (perpetuated through systems use or enactment), and highlighted the way this may depend on the degree to which a system or IT artefact imposes a view or Knowledge on users. Now I turn to how such agency may vary for particular Groups involved and over time by reflecting on the cases observed.

For document controllers within InfraDig, the eB system was encountered as a given, whereas for the asset Data team, it represented a more malleable artefact. This reflects the latter's ability to influence its design, within the boundaries of what was configurable. The original eB system designers were removed and yet powerful in having originally set these boundaries, and indeed still powerful in their ability to change them, through subsequent releases of the software, that may in turn be enacted through implementation and use. The

ultimate users, the Operating company practitioners, were relatively removed from the design process, other than through limited, abstract design requirement discussions. For them, the final system and Data artefact delivered may well be encountered as a given. On the other hand, they have the ability to choose whether, to what extent and how to integrate the artefact provided within their practice and reified processes and systems.

This illustrates various roles and power relationships at play between authorial, interpretative or enactive groups and the IT artefacts they interact with. This may well reflect their relative agency in any resulting nexus. It also points to how these may shift, for example over time, as designer or Data capture Groups disappear from the 'scene of the immediate action' or arena of practitioner performance in view. Orlikowski's (1992) example of the view of experienced users (pre-dating their system introduction) varying significantly from subsequent, new users is instructive. For earlier, experienced users, the system represents a particular set of choices and trade-offs (amongst several options), which they don't necessarily feel bound or constrained by. By contrast, for new users of this system, questions of constraints and options do not arise, the system is simply accepted as a given. In this instance, the system actant can be thought of as having far stronger relative power or agency in relation to them than in relation to experienced practitioners.

These examples highlight the potential impact of the experience of particular practitioners (or Practitioner Groups) on the relative agency and power attributable to material and human actants in relation to particular system and Data components. They also shed new light on the importance of design involvement.

The value of a Practitioner Group view of Authorship over time

As illustrated by the case observations, Knowledge about the Phenomena, Data use and its application typically resides in the ultimate Practitioner Groups engaged with the Phenomena and related practices. Knowledge about optimal Data organisation, processing and storage typically resides within IS Practitioner Groups, for example in programmers, architects, etc. A limited

number of Boundary Spanners may also exist, who have some Knowledge and experience of both Knowledge domains, e.g. expert users, business analysts or Data analysts.

My research illustrates how economies of meaning, their ownership and negotiation are important aspects of how Practitioner Groups make their Knowledge and learning explicit, and allow this to evolve. This was particularly apparent in OtherCouncil's development of a deprivation index, which gradually stabilised in terms of dimensions and measurement over a period. It was also apparent in the familiarity and standardisation inherent within Acorn Data. This represented a Marketing specialist domain area. Similarly, the terminology used within InfraDig reflected considerable institutionalised engineering Knowledge. Participants noted the need to develop and agree even greater standardisation between different Practitioner Groups, reflected in BIM related industry initiatives.

Closely related, my research also shows how different economies of meaning and related reifications may represent Barriers to Data Engagement for other Groups. Certainly, this was something I experienced on entering both research settings, being unfamiliar with Acorn Data and Marketing at GoCouncil, and similarly unfamiliar with much engineering construction terminology at InfraDig – both represented learning curves, which took time. However, my introduction of CoP concepts and thinking within InfraDig also showed how introducing new terms and concepts facilitated their being able to better articulate certain issues, such as their communication issues between Groups.

The emergence, management and evolution of such economies of meaning should clearly be an important part of theorising IT artefacts or systems. These essentially represent reifications of such Knowledge and meanings, and can represent important Boundary Artefacts between different Practitioner Groups.

Losing sight of group authorship may exacerbate the tendency highlighted towards taking the IT artefact presented as a given and unalterable. The tendency to take the IT artefact as a given represents an increasing investment and institutionalisation of particular economies of meaning, which may lead to

increased inertia or anchoring of thinking. Recognising a clear link between authorship and institutionalisation further underlines the importance of distinguishing and separating Data and algorithmic elements argued for earlier.

Firstly, it highlights the different authors and activities involved over time. For example, system designers of Data elements, structures and relationships and algorithmic elements are central at the design stage, while Data capture staff may have an ongoing authorial role. Data capture staff could be further differentiated between those Groups capturing detailed 'transactional' Data, versus those who may make changes to master data and parameters.

Secondly, it brings into focus the question of what is left tacit or implicit by such authors, or the tacit precepts commonly assumed to be understood. This can be inherent or implicit in structural and relational choices reified, e.g. Data category options. Thirdly, different authorial Groups may have greater opportunity for reinterpreting during use than others, which may result in related elements evolving or changing at different rates. For example, structural Data and algorithmic elements may remain relatively stable (or rigid and inflexible) compared to the potential for creating new Data entries and categories, or reinterpreting available categories during Data capture as new situations or possibilities arise.

Such reinterpretive acts may also have different levels of scope and impact. For example, changes to transaction Data during Data capture may have a limited or local impact, while changes to master data parameters may have a more significant or pervasive impact. This may also be closely associated with the relative power and importance of particular author Groups, often reflected in differentiated access rights to update Data.

Accounting for Data's role as a Boundary Artefact in some contexts

Closely linked to recognising the importance of group authorship, the case findings outlined in section 7.6.3, indicate how Data acts as a Boundary Artefact or infrastructure between different Practitioner Groups involved. They address shared economies of meaning that are relevant to multiple Groups.

Star (2010) draws a distinction between Boundary Objects, which facilitate sharing and use without consensus on meaning, versus Boundary Infrastructures, which establish more formal shared meanings. Indeed, she calls for research on how Boundary Objects may develop into Infrastructures. For instance, the same Data may start as a Boundary Object and gradually become a Boundary Infrastructure over time, with greater shared understanding of related field and category meanings, or it may not, depending on how it is used and Engaged with. Establishing related meta-Data, Data mappings, Data models and dictionaries can also be seen as attempts to aid a transition towards shared meaning and such infrastructure.

To theorise Data in relation to generating insight, its ability to facilitate productive conversations (Tsoukas: 2009) also seems important to investigate further. His work recognises the facilitation role that Boundary Artefacts can play between different Groups, which would be useful to investigate in relation to Data as a specific kind of Boundary Artefact. It would be interesting to contrast its use in different practitioner or disciplinary domains or across them. Such an effort could build on work by Edwards et al. (2011) in relation to interdisciplinary Data and Knowledge sharing, and on work about how artefacts aid collective sensemaking more generally (Stigliani & Ravasi: 2012). Such Data use may also differ between exploratory Data Engagement and more exploitative Data Use.

Some preliminary implications for additional MSS model constructs

The argument for disentangling Data from the wider MSS artefact, clearly prompts the inclusion of a Data construct in the MSS model. In addition, the following related constructs may also be useful to include:

- Authorship (already discussed above)
- Sensing and Scanning activity for new Data and Knowledge
- a Data Quality/Validity construct

These reasoning for adding the last two additional constructs noted above are briefly outlined below in turn. Given the need for considerable further research

and theorising about Data, these suggested additions are obviously and necessarily provisional.

Sensing and Scanning for new Data and Knowledge

Sensitivity to different Knowledge domain areas broadens our related notion of Sensing (Kettinger & Marchand: 2011) as an information management practice. This is extended beyond the four external focus areas currently identified, to encompass a much broader set of possible, relevant Knowledge domains. These are likely to depend on the nature of the Phenomena, as well as Data and practice improvements in view. At the same time, this provides an organising principle for such scanning, which clearly links back to particular MSS components and related Data or domain Knowledge encoded.

The notion of sensing can be further extended, to encompass the ongoing scanning and evaluation activity in relation to new Data sources and services becoming available, e.g. public Data in terms of open.gov initiatives in the UK and the like. This seems very similar to the existing notion and related approaches for technology horizon scanning and emerging technology evaluation used to inform IS strategy formulation. This goes far beyond considering new external variables that may be relevant to ensure a good MSS Problem Space Match and points to a need for a broader concept of scanning to be encorporated in an MSS model and for making it explicit.

A Data Quality/Validity construct

Including a separate Data construct in turn prompts us to reconsider the notion of a single measure of MSS Quality. It argues for a separate, distinct Data Quality construct. Indeed, Data Quality has long and widely been recognised as an important consideration for Data Use or Engagement. This is often referred to in terms of the level of confidence or trust placed in Data and related analysis (Davenport 2014, Davenport et al.: 2010, Marchand et al.: 2001).

Within the MSS model, the MSS Quality construct itself is recognised as under-theorised, with related factors giving rise to it not being identified. The provisional definition the authors provide is in terms of the MSS providing

effective support to decision-making and a good MSS Problem Space Match. A separate Data Quality construct certainly seems a step forward, pointing to further elements of MSS quality too, for example in relation to the quality or maturity of the Knowledge encoded. Authorship and shifting interpretation seems an important consideration for any proposed Data Quality construct, pointing to the likelihood of different quality assessments for different user Practitioner Groups. These groups are likely to use different criteria and tolerances or thresholds, depending on their proposed Purpose and context.

To the extent different Practitioner Groups use Data in a similar way, Data Quality assessments might pragmatically be grouped. This may be more likely for stable Data requirements typical of Data exploitation contexts than for exploratory contexts. Data Quality assessments should ideally be performed by the Practitioner Groups using the Data, even if the Data are captured by other Groups.

This underlines the need for communication and coordination mechanisms between these Groups, to agree quality standards to meet different requirements. Such mechanisms should also address related costs of Data collection, with related cost sharing by user departments where necessary, to ensure alignment and adequate Data quality. This seems likely to be an important component of related Data governance arrangements. However, if Practitioner Groups' use or inquiry varies over time, or in relation to different inquiries, the usefulness of such Group quality measures may be limited. Instead, adopting a context or inquiry specific assessment may then be more appropriate.

The value of research validity concepts emerges again as a useful avenue for theorising the construct and for related practice. This would tie Data quality to the conclusions being drawn from the Data analysis performed, and any claims being made. A broad concept of validity also has the advantage of not only recognising the importance of narrow, technical *statistical* considerations when drawing inferences from quantitative Data analysis (Davenport: 2010), instead, it can encompass both qualitative and quantitive Data and related analysis.

Linking the quality assessment to an explicit inquiry logic can help refine both the logic as well as the Data requirement. This may be especially useful early during a Data initative, much as a pilot study might be in a research setting, to validate the proposed research method.

Figure 8-12 visually summarises the ideas introduced above, in terms of additional considerations or constructs to reflect in a model in relation to Data.

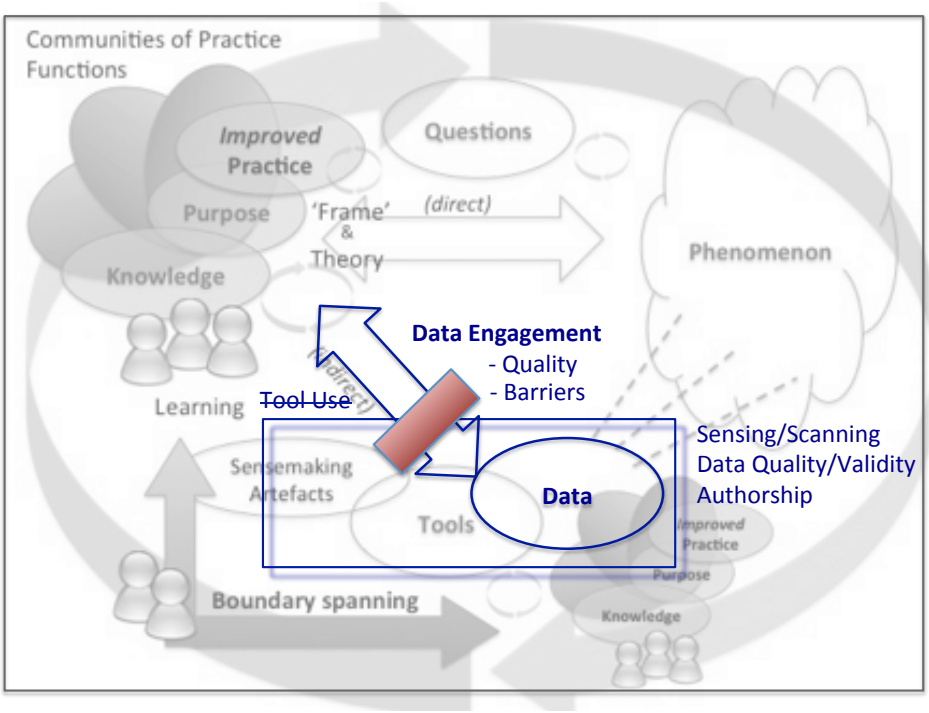


Figure 8-12 Important Data attributes and activities in relation to pursuing insight

8.1.5 Address different practitioner groups and Knowledge domains

Several of the proposed improvements to theorising already outlined are predicated on a clear view of Practitioner Groups involved. The argument to disentangle Data has further highlighted how encoded domain Knowledge is often related to particular Practitioner Groups, drawing attention to their authorial role in developing and maintaining encoded economies of meaning.

However, in the MSS model, human participants are simply characterised as users (advanced analysts and end-users), or management sponsors and supporters of the MSS (the central focus). Their Commitment, level of MSS Use and MSS Knowledge are presented as related elements. Meanwhile, MSS

designers are not reflected in the model at all, an important omission given their central role and significant influence.

Locating Data initiatives or MSS within and across the Practitioner Group contexts involved is therefore important. Doing so facilitates an understanding of different Groups' associated Knowledge and learning, as well as particular contextual factors, e.g. internal politics and relative power relationships between Groups involved in the initiatives. As illustrated in my research observations, these may vary, and Group level assessment may be more feasible than doing so at an individual level of analysis. Some of the related Group aspects or ideas that might be addressed in an improved MSS model are outlined in turn below.

Developing local domain Knowledge

As already highlighted above, the state of particular Practitioner Groups' local domain Knowledge and experience is likely to be important when assessing the scale of ambition related to a particular Data initiative. This is especially true for domain Knowledge that relates to the ultimate Phenomena and related practice. However, it may also be relevant for any domain Knowledge that is critical to enabling Data gathering and Data analysis. For example, it could also relate to new or unfamiliar technologies or Tools that will be used to collect or organise the required Data. The current state and rate of improvement or progress envisaged for domain Knowledge areas, as well as the related scale of learning ambition this represents, will be important considerations. This is crucial to correctly estimate project Complexity, effort, timescales and related risks to realising envisaged benefits.

The above points to the value of explicitly recognising the different Knowledge domains relevant to the overall initiative. This could possibly be reflected in a distinct Domain Knowledge Construct within the model or addressed as a dimension of the Theory Quality/Maturity construct suggested earlier. Associated measures of current domain Knowledge maturity, and aspirational ambition, while unavoidably subjective may still be pragmatically useful to gauge the relative effort and risk associated with different projects, especially when evaluating or assessing their progress at a portfolio level.

Wider IMP, Data orientation and related capabilities may also vary by Practitioner Group. These represent further considerations when assessing their domain Knowledge. Wenger (1998) recognises that different Practitioner Groups may have a different balance between how much and what Knowledge is made explicit about their practice and what is left tacit. This may be an important part of considering their likely Data orientation, the degree to which their Knowledge will be straightforward to encode, as well as anticipating likely resistance or Barriers to Data engagement and MSS or Tool Use.

Practitioner Group capabilities and activities to acquire and integrate new Knowledge, investigate theories, test or apply them, and keep up with domain developments, are likely to be similarly important. In addition to the concept of *sensing* already introduced (Kettinger & Marchand: 2011, Marchand et al.: 2001), the concept of absorptive capacity seems highly relevant (Cohen & Levinthal: 1990). While these are organisational level concepts, they may be worth considering at a project level for particular Practitioner Groups too.

Focusing on Alignment for emergent learning

It seems natural that different Practitioner Groups are focused on different activities and objectives. Their concerns are likely to be focused on their particular, instititutionalised domain Knowledge and central Phenomena. At InfraDig Groups specialised in particular activities: Data collection, MSS or Tool design and configuration. These groups contributed specific domain Knowledge, as described in section 7.4 and illustrated in Figure 7-3. My research observations illustrate how this in turn led to institutionalised or bounded practitioner framing or design thinking, e.g. the Data elements or dimensions thought relevant and application Data schema adopted.

The question of Alignment between Practitioner Groups around the ultimate inquiry and benefit objectives therefore seems an important consideration for theorising such initiatives and practically managing them. Alignment may also mitigate the risk of becoming anchored and bounded in the Data being produced, rather than the broader objectives. Indeed, Wenger (1998) emphasises the importance of Alignment activities to promote emergent

learning and that they should be included in learning design architectures. My research observations, particularly for InfraDig, highlight a lack of consensus between participants and Groups as to the overall Purpose of the Data initiative, as well as an absence of the ultimate maintenance Practitioner Group. As a result, learning about maintenance was certainly not a recognised, shared focus and was effectively deferred.

Where multiple Practitioner Groups are involved, it also seems important to be clear about the ultimate or primary Practitioner *User* Group. By contrast other Groups involved are secondary, though important enablers of particular elements of the overall effort. Enabling Groups' understanding of the wider or ultimate learning objectives, and clarity about their enabling role, may avoid the silo or local thinking mentioned by many participants across both cases.

Based on the above, considering Group Alignment and its assessment is likely to be important at different levels:

- within particular Practitioner Groups, in terms of how particular economies of meaning emerge, are negotiated and come to dominate;
- between Practitioner Groups within a particular organisation; and
- at a wider level where Practitioner Groups span organisations or sectors.

Alignment challenges are also likely to reflect different Groups' power or influence over other Groups. Various constraints on Practitioner Groups' influence over other groups were noted across both cases studied. Different strategies were also employed by leading Practitioner Groups to address Alignment. This surfaced particularly prominently in the InfraDig case given the number of Practitioner Groups and organisations involved, as well as the contractual nature of many of these. Section 7.4.4 describes a complex web of tactics and mechanisms employed to achieve alignment. These combined line authority over some Groups with internal influencing of others, and contractual oversight mechanisms with third party Contractors. As highlighted in section 7.4.1 above in relation to Contractors, conflicts of interest can arise in such situations. It also illustrated how InfraDig had far greater leverage over Contractors than Operators, the latter being essentially their clients.

While GoCouncil's organisation was more straightforward, power and control over resources are nevertheless important. At one level, not explicitly acknowledged, the very reason for Marketing becoming involved in the corporate business planning project, was to remind the organisation of their vital value. In the cost-conscious climate they may well otherwise be seen as an overhead that should be trimmed. Nevertheless they are still subtly challenged by senior board members in relation to how much market insight is really required to inform business models.

Implications for MSS model constructs

The above sections point to the need to recognise when different Practitioner Groups are involved, as a particular dimension of MSS Complexity. An associated, specific and new measure of Alignment also seems sensible, distinct from MSS Problem Space Match. This should address both the degree of consensus as to scope, objectives and priorities, and the overall Alignment of activities and Data to the ultimate agreed Phenomenon of interest in the inquiry logic pursued. Relative Practitioner Group Influence is important but it is not clear how best to reflect it within the MSS model. While it might most easily be included as a further dimension of MSS Complexity, this may not provide enough prominence. It may instead warrant explicit inclusion, especially where a number of different Practitioner Groups are involved.

Both cases also illustrate how the leading Practitioner Group's influence over Data collection can be different from their influence over subsequent use, practice improvements and benefit realisation if these are pursued by different Practitioner Groups. This points to the potential value in considering the question of power and influence separately in relation to the inquiry effort and the practitioner benefit realisation effort and related logic elements. A separate construct may make it easier to identify and link or map this factor to other constructs in the MSS model. The wider management literature on power may also be a useful source of further theorising constructs or dimensions for potential inclusion. Reflecting the discussion above, an additional concept of Alignment is shown in Figure 8-13. While this is positioned at the overlap

between Inquiry and Benefit Spaces, Aligning various Groups within either Space or across Spaces is also likely to be important. The question of domain Knowledge maturity is also reflected in terms of the earlier Theory Maturity/Quality construct introduced in earlier sections.

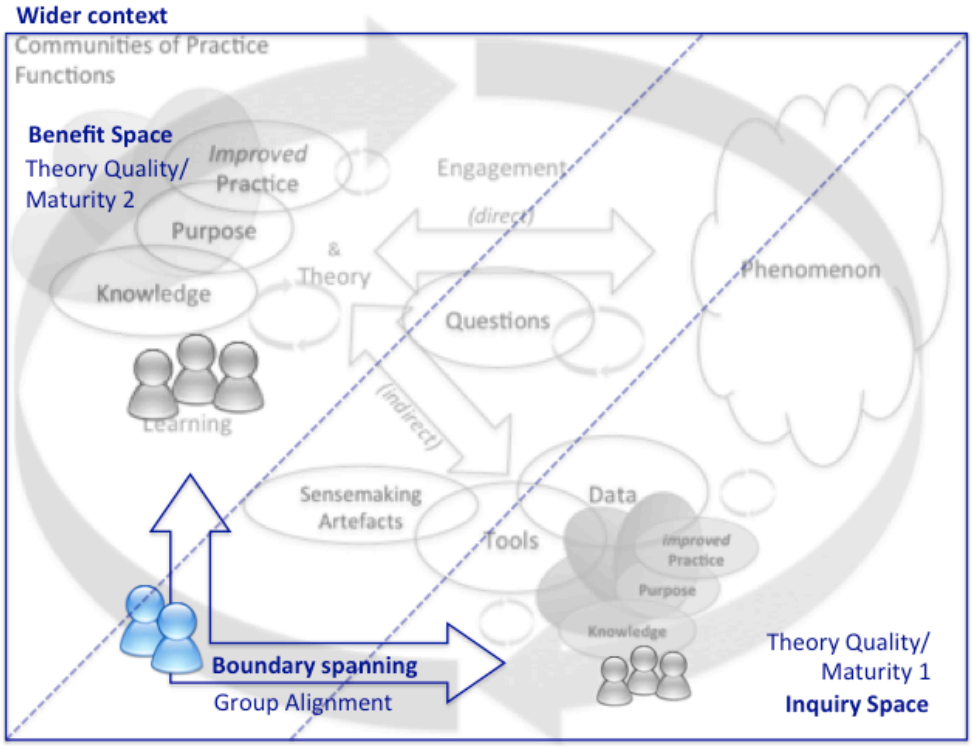


Figure 8-13 Aligning Practitioner Groups across Inquiry and Benefit efforts

Given the various mechanisms and strategies employed to achieve Alignment and coordination within such initiatives, the Boundary Spanning element of the explanatory framework is also introduced, which is examined next.

The Importance of Boundary Spanning Activities

Where Data initiatives span Practitioner Groups, as highlighted in InfraDig, it seems likely that the scope and quality of Boundary Spanning activities and related Boundary Spanners will be important to enhance Alignment and related emergent learning. The proposed importance of Boundary Spanning is consistent with findings about the broader importance of wide extra-project networks for success in complex, *Knowledge-intensive* initiatives (Cummings & Pletcher: 2011). The importance of project team members' personal networks is emphasised by their research.

Boundary Spanners may be particularly important, given their likely broad networks and experience of different domains. They are likely to promote Alignment, as well as Engagement and innovation across particular Group boundaries or edges. For example, Boundary Spanners are more likely to keep abreast of wider developments, especially in a different domain area, and then be able to help local practitioners see the relevance for their practice or domain area (Cohen & Levinthal: 1990). They may also more readily challenge institutionalised thinking. Work on the importance of boundary 'shaking' for achieving strategic change (Balogun et al.: 2005) may be usful for further theorising in this area.

This seems to be an area worthy of further research to improve our theorising about multi-disciplinary data initiatives that span Practitioner Groups. Several potential lines of further inquiry are outlined below:

- Project manager, business analyst and MSS designer Practitioner Groups drawn on for Data initiatives deserve particular attention, as they tend to play pivotal roles in shaping initiatives, Data and MSS design. Such research might encompass the extent to which they focus on and support or undertake Boundary Spanning activities. It could also focus on the extent to which their particular prior experience of the ultimate practitioner domain versus other domains proves useful.
- Further research on Boundary Spanning more broadly could look at the effectiveness of particular Boundary Spanning activities and strategies, as well as Boundary Spanners, to examine their relative effectiveness for different contexts of practitioner power relationships and dynamics, as well as considering wider corporate cultural factors. Such research should also take account of variations in the scale of organisations and their project efforts, which may represent important dimensions of MSS Complexity.
- As already argued earlier in relation to embracing an innovation framing for exploratory Data initiatives, sensing, exploratory and Boundary Spanning activities may require slack resources (Salge and Vera: 2013).

Boundary Spanning may also not be considered core or a priority for particular Practitioner Groups, particularly when under time or other constraints. Indeed, in common with research activity more generally, these activities trade immediately relevant and urgent benefits for the promise of future, uncertain benefits. Specifically planning slack resources for exploratory initiatives may therefore allow Practitioner Groups time for exploration and development activity. Such activity may also be more common or possible in profitable or resource rich organisations or sectors.

These ideas all indicate a rich vein of further inquiry and research to better understand and theorise emergent learning in these settings.

8.1.6 In summary

The above argument demonstrates how a focus on the underpinning logics being enacted or pursued within Data projects, helps reframe the MSS model. The richer model that emerges more clearly addresses the wider practitioner pursuit of insight by engaging with Data to help illuminate a Phenomenon, in order to inform and improve practice and achieve desired outcomes.

A visual summary of the various related MSS enhancements proposed is presented in Figure 8-14, reflecting them in relation to the explanatory framework introduced in Chapter 7.

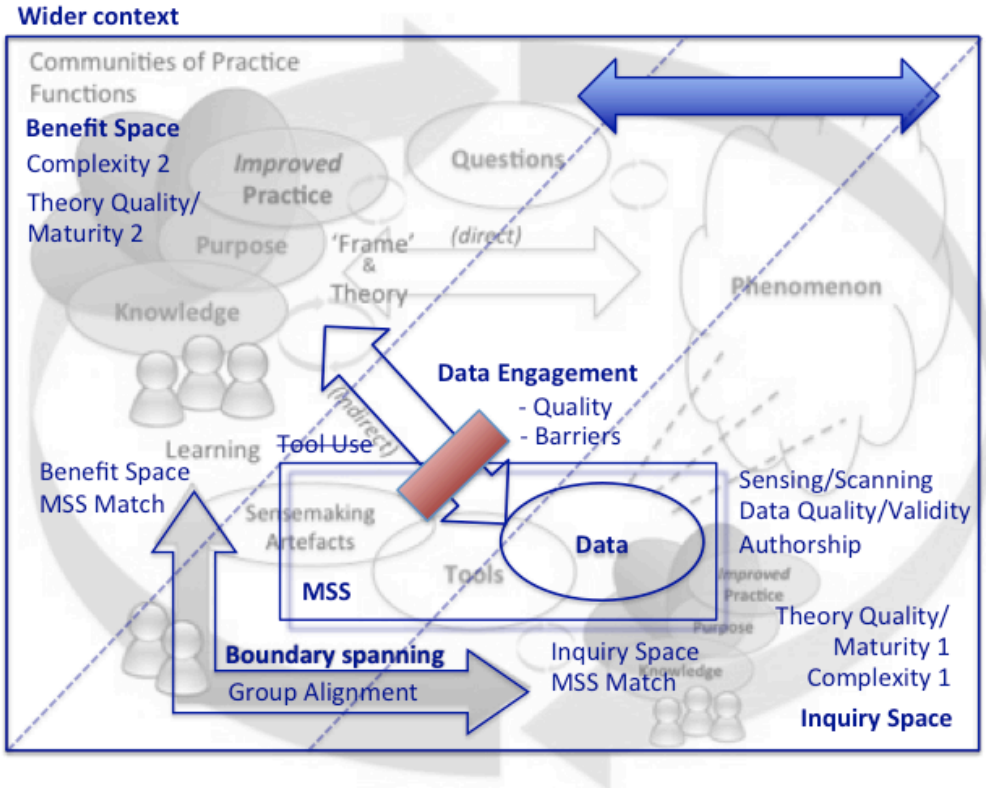


Figure 8-14 A visual summary of proposed MSS enhancements

The framework is reflected in the background with the following MSS improvements discussed above highlighted or foregrounded:

- The Problem Space is reframed, instead introducing overlapping and interdependent Inquiry and Benefit realisation Spaces. The associated MSS Problem Space Complexity construct is split into two to reflect this. A new Theory Quality/Maturity construct is also introduced, in relation to the underpinning inquiry and benefit realisation logics at work, and related domain Knowledge being used and related learning pursued.
- The MSS itself is contextualised or mapped to Tool, Data and Sensemaking elements of the framework. The related MSS Problem Space Match construct is split in two, relating these to the new Spaces or domains introduced.
- Data Engagement is highlighted as central to generating insight from Data, rather than MSS/Tool Use (shown as crossed out). This prompts the introduction of two new related constructs as important: the Quality of

Data Engagement, as well as recognising various Barriers to Data Engagement.

- Most significantly, Data are recognised as a specific construct, disentangled from the wider MSS artefact and algorithmic Tool elements. This prompts the introduction of several new related constructs as important considerations: Sensing/Scanning for new Data that may be relevant, Data Quality/Validity in relation to the Phenomena in view, and recognition of Data Authorship their and related social construction.
- Finally, the Boundary Spanning nature of Data initiatives is also recognised, introducing it as a construct and Alignment as a consideration, particularly where different Practitioner Groups are involved.

8.2 Adopting a CoP lens to study Data initiatives

Having outlined various areas for improving my IS theory and models for Data initiatives and related MSS above, this section turns to questions of method or techniques. This contribution was not a particular focus encompassed in the original research question posed. Rather it emerged when reflecting on my method and the sensitising theoretical lens used to make sense of my cases.

The section starts by arguing that adopting Wenger's (1998) Commmunities of Practice (CoP) framework as a lens and framework has value for both practitioners undertaking Data initiatives, as well as for researching such initiatives. It then proposes group or community mapping as a useful technique to capture, understand and present a project's different practitioner contexts. The concept of boundaries is also examined more closely, arguing for the importance of relating these to functional and organisational structure, to identify where they reinforce each other or introduce a tension between practice and organisational structure and related boundaries.

8.2.1 Practitioner and research value

For researchers, using the CoP framework proves useful in two main ways. The first is that it helps to bound or situate the research and Phenomenon within a particular practice setting. While this is similar to the idea of focusing on 'site-shifting' adopted by Huang et al. (2014) within the strategy-as-practice literature, Wenger's (1998) framework allows us to manageably examine both group and individual level aspects of the Phenomenon, bringing boundaries, related spanning activity, artefacts and agents involved clearly into view.

The characteristics of site-shifting that Huang et al. (2014) highlight which promote ambidexterity, could also apply to particular practitioner communities. In turn, practitioner learning and related negotiation of economies of meaning, may provide one potential mechanism for how site-shifting might occur. This raises interesting, further research questions: Does site-shifting give rise to learning, learning give rise to site-shifting, or indeed, could both be at work and what might influence this?

The second major benefit of adopting CoP, illustrated in Chapter 7, is the framework's conceptual integration of learning, Knowledge, artefacts and Tools, providing very useful concepts and terminology with which to examine, analyse and then explain what is observed. Examples include concepts such as economies of meaning, Boundary Artefacts and Spanners and their importance for productive dialogue (Stigliani & Ravasi: 2012, Tsoukas: 2009). This sensitised and alerted me to relevant related cues during Data collection, and prompted reflection during Data analysis, generating new insights. In particular, this brought practitioner Data use, as well as Data's sociomaterial nature into sharper focus, both as a boundary artefact and as a codification or reification of practitioner Knowledge about a Phenomenon. Importantly, it helped distinguish Data from the closely associated algorithmic elements within an IT system or artefact, allowing Data to be examined more closely in their own right. This focus on *Data* clearly allows us to build on and extend earlier work by Orlikowski (1992) on duality and sociomateriality (Orlikowski: 2006) of

technology, while sharing an emphasis and focus on situated practitioner *knowing* and technology use (Orlikowski: 2002, 2000).

Adopting a CoP lens also proves useful for practitioners. While InfraDig participants identified 'changing mindsets' as a major challenge, they found it difficult to articulate the challenge more specifically. Introducing them to the CoP framework made it easier to 'frame' the problem, identify which Practitioner Groups were involved, what learning was envisaged, and where to focus attention to improve collaborative engagement (e.g. forums being required, lack of engagement, etc.).

These observations accord with and complement work on strategy blindness and cognitive entrenchment (Arvidsson et al.: 2014), which had a primary emphasis on practitioner *learning* rather than a focus on *changing* practice, although these activities seem likely to be closely interrelated. In the InfraDig case, this made the challenge more manageable for practitioners to bound, and shifted the emphasis from broad communication to considering how to facilitate more specific engagement. It also offered a common language with which to unpack and discuss challenges and proposed interventions.

8.2.2 Mapping communities

Visually mapping CoP involved in a Data initiative, as described in section 5.3, provides practitioners and researchers with a useful cartographic technique. It can be used to understand and analyse the various practitioner groups involved, their reified artefacts, formal and informal forums for negotiating meaning, as well as boundary spanning activities and boundary artefacts.

For practitioners, this can also provide a useful point of reference with which to consider broader, multifaceted interventions and proposed practice changes, recognising the scale of the change, addressing tacit and explicit elements for different Practitioner Groups, better tailoring the framing of an initiative for participant groups to encourage participation and engagement. In InfraDig's case this was only partially realised with the mixed adoption and initial

resistance from the head of asset Data. This may underline the need for such artefacts to be jointly created with participants in order to be accepted by them.

This approach represents a form of Data visualisation that may be particularly intuitive for researchers for certain aspects of Data exploration and analysis. For example, this could be further developed through the use of iconography and colour coding, facilitating the creation of related 'heatmaps'. These approaches can be directed to address a variety of different assessments and questions. One such example might be to assess the degree of reification of practice and related Knowledge within different Groups mapped (i.e. the balance of tacit and explicit core Knowledge related to the practice area or group). Another may be to assess different, relative levels of Knowledge about a Phenomenon assessed, or degrees of longstanding, unchanged institutionalised practice, versus groups evidencing considerable innovation and change in their practices and related reifications.

The technique could provide an interesting and pragmatic complementary or alternative technique to network analysis, which typically captures Data at the level of the individual, rather than groups, and usually uses communication as its focus or starting point (Cross et al.: 2006). By contrast, when used in conjunction with a flexible tool such as Prezi, a great deal of detail can be captured about Groups or communities, allowing for easy Data Engagement at different levels of detail, e.g. zooming in and out and alternating between overview analysis and engaging with detailed artefacts. The approach may also complement more traditional systems analysis techniques such as high-level Data flow diagramming, providing a way of locating such logic within a broader human setting. The approach may also help address the recognised challenge of how to address sociomaterial practices effectively in text and writing (Cecez-Kecmanovic et al.: 2014), perhaps in conjunction with specific *zooming in* and *zooming out* techniques used within the sociomaterial research context (Nicolini: 2009).

These ideas represent a rich vein for future practical investigation and experimentation. An accumulated body of such maps collected over time, may

also provide a potential Data source for higher-level topologically oriented comparison and analysis across different Data initiatives and settings mapped. This might be thought of in a similar vein to network topologies, although in a more interpretive, geographical sense, rather than a mathematical sense.

8.2.3 Practice Boundaries, Edges and Borders

Wenger's (1998) concept of CoP emphasises people united by a common practice they engage in, and provides useful ideas about practice Boundaries and peripheries. He identifies different types of boundary encounter (one-on-one, immersion and delegations) between different practice areas, all of which were noted in the cases observed, although immersion was not noted in relation to Data practitioners within InfraDig. No Operator Data practitioners were immersed in the Operations handover team or elsewhere, nor were InfraDig Data analysts embedded within Operators. Wenger also goes on to identify Boundary practices, overlaps in practice and peripheral engagement as useful forms of connection between practice communities.

The ideas Wenger (1998) contributes in relation to Boundary Objects or artefacts in such learning settings, proved particularly useful during my research. While he bases his ideas on Star's (2010) Boundary Object concept, he highlights their features of modularity, abstraction, accommodation (of different activities) and standardisation. Indeed, this idea is also explored further in work by Tsoukas (2009) in the context of facilitating *productive dialogue* between different practitioner groups in connection with creating new Knowledge.

However, the use of the term 'Boundary' in connection with Boundary Objects is worth clarifying. Star (2010) recognises that its normal use may imply a periphery or edge, while she wishes to imply a 'shared space' between practice areas, which was not clearly within one or other practice area. She also argues for a distinction between Boundary Objects, as organic structures to enable different practice areas to collaborate without consensus, and more formal, longer lasting Boundary Infrastructures.

While Wenger (1998) recognises Boundaries between practices as important, and acknowledges that institutional Boundaries may or may not coincide, it is not clear if he refers to shared peripheral spaces or practice *Edges*, though the latter seems more likely. This points to the need for clearer terminology to describe and distinguish shared peripheral spaces from practice Edges. This is echoed in criticism of the CoP cited in the literature review chapter (Hughes et al.: 2007). More importantly, Wenger's (1998) framework work does not specifically address the *tensions* he acknowledges may exist between such practice Edges or peripheries and more formal institutionalised Edges. More formal organisational Edges or demarcations are differentiated from practice edges and termed *Borders* in this discussion to avoid confusion.

As anticipated by Wenger, my research findings highlight how Data initiatives often span both practice Edges and internal and external organisational demarcations or borders. For example, in the case of InfraDig, the broad engineering community spans various organisations, with many shared economies of meaning, even while particular local practice foci may be more differentiated. This introduces various tensions, e.g. in relation to practitioner identity, negotiating or coordinating different economies of meaning, resource allocation, relative and perceived power differences, alignment and coordination of purpose and related incentives, as well as activities and shared or co-produced artefacts, e.g. design 'as-built' documentation.

In particular, the findings in section 7.7.4 above highlight that some practice Edges or peripheries are 'harder', or less permeable than others, especially where they coincided with functional (internal) organisational Borders. This quality of 'hardness' or 'lack of permeability' was noted to an even greater extent when practice Edges coincided with contractual and external organisational Borders. In those cases, this quality of 'hardness' or lack of permeability served to restrict or constrain communication, collaboration, peripheral boundary-spanning activity or engagement.

This may in turn have implications for the emergence and occurrence of more formal or standardised boundary *infrastructures* rather than Boundary Objects

(Star: 2010). This prompts whether or not such formal *infrastructures* are based on or developed from informal Boundary Object precursors in all cases, or if the starting point and development depend on the permeability of the pre-existing peripheries and level of informal engagement they enable.

Importantly, these tensions and 'tribal' identifications are observed to *transfer* to related Boundary Objects or infrastructures such as Data (see section 7.6.3 above), especially if different economies of meaning are adopted. This may cause further tensions and Barriers to Engagement with Boundary Artefacts, especially formal *infrastructures*, perceived as 'owned' by others rather than co-produced or locally 'owned' by the local group or CoP. This effect seemed lessened or mediated by physical co-location of practice areas (e.g. in the instance of document controllers on site from both Contractors and InfraDig), and to a lesser extent by shared membership of a broader common practice community or professional identity (in terms of a common economy of meaning or language). More formal Boundary Artefacts or Infrastructures and related contractual arrangements clearly serve to institutionalise or reify practice, potentially introducing rigidity or standardisation to related interactions and Data sharing.

As illustrated in my research observations, framing Data as a Boundary Artefact or Infrastructure provides a very interesting line of inquiry for further research about Data, intrinsically recognising its socially constructed nature and its fundamental link to and bridging of different practice contexts. Wenger's (1998) argument that reifications may not travel far on their own, without practitioners to assist with appropriate interpretation in relation to the original action context it was created in, seems particularly pertinent to explore in relation to Data. Such research would also address a wider call for more empirical work on Boundary Artefacts and more formal Infrastructures (Star: 2010) as well, and the role of Boundary Artefacts in helping to facilitate new Knowledge creation between different Practitioner Groups (Tsoukas: 2009, Hughes et al.: 2007). Indeed, it offers a further synergistic starting point for research collaboration between Data and Knowledge Management disciplines.

The discussion also underlines the need to extend the CoP Boundary and Edge related concepts and theorising, as argued for by Hughes et al. (2007). This should explicitly address tensions that arise from interactions between formal organisational structural Borders, practice Edges and practice Boundaries. This would certainly make the CoP framework more useful to address communities that span such Borders, as well as cross-disciplinary Data, learning and innovation initiatives.

This need seems familiar and closely related to issues of identity, coordination and alignment within matrix organisational structures. Therefore, literature and research concerned with promoting collaboration across organisational units and between organisations, seems relevant and may prove a useful source and starting point for further theorising. In turn, using an extended, group level Community of Practice lens may also provide a promising line of research inquiry to address issues and research questions in that domain.

8.3 Summary Contribution Table

While the main contribution of my ethnographic research can be argued to be the thick case descriptions and related explanatory framework outlined in earlier chapters, this discussion section has sought to relate these to existing theory. This has primarily focused on potential areas for refining or extending the existing MSS model within IS, highlighting several (social or human) elements that seemed important in the cases observed. It also highlights the need for better conceptualising and theorising Data as an IT artefact, drawing on socio-materiality.

A secondary area of contribution relates to the value of use of the Communities of Practice framework and lens for studying data initiatives, especially the approach of visually mapping out practitioner communities involved, their knowledge artefacts, areas of engagement and boundary spanning, etc. My research also lends weight to calls within this area of research to better define boundaries or edges and borders between such communities, offering one or two useful ideas to take this forward.

The table below seeks to summarise the proposed areas of theory refinement or extension contributed by my research and synthesis of related theory considered. It highlights how the observations and related discussion either address particular gaps identified and/or point to future research required to address and further refine our theory for such initiatives and related systems.

Table 8-1 Summary of Observations, Contributions and Implications

Observations	Relevant Related Research	Gap/Deficiency	Contribution/ Claim/ Implication
Reframing and contextualising IS models for data initiatives			
Inquiry and benefit logics pursued are not distinguished from each other and are left implicit by case participants and in the MSS model	**MSS Model Framing** MSS Model (Clark, Jones & Armstrong: 2007) versus Sensemaking and framing challenges (Weick: 1995, Schon: 1991) Espoused theory versus theory-in-use and related logics aimed at understanding versus application (Argyrus & Schon: 1974, 1978, March: 1991) Importance of context generally (Lee et al.: 2015, Checkland & Holwell: 1998) and to encourage ambidexterity (Huang et al.: 2014, Turner, et al.: 2013), for decision-making and related pragmatic data use (Snowden & Boone: 2007, March: 1987, Davis & Olson: 1984, Mintzberg: 1975, Lindblom: 1959) Bricolage and material approaches adopted by practitioners (Duymedjian & Ruling: 2010, Stigliani & Ravasi: 2012) Innovation versus problem framing (Arvidsson et al.: 2014, Leonard & Swap: 2006, Eisner: 2003, Davis & Olson: 1984)	MSS focus on system and tool use rather than practioner inquiry, sensemaking, learning and action The IS literature's focus on improved decision-making as the natural benefit flowing from MSS is problematized as too narrow and abstract to be useful for practitioners The practitioner context is simply framed as a problem-space, rather than linked to practitioner challenges, central phenomena and related knowledge domains, or allowing for an exploratory and innovation focus A unitary view of the MSS does not acknowledge pragmatic practitioner use of a variety of resources and systems to achieve their objectives	A broader explanatory model is synthesised, from a Practitioner Group starting point, repositioning Tool Use and related MSS model elements. **See also Practioner Group observation/contribution** The MSS Problem Space construct is reframed as an Inquiry Space, removing a potential biase towards problem-solving rather than innovation or exploratory inquiry A separate, though overlapping and interdependent, benefit realisation space is distinguished, focused on acting on improved information, knowledge and theories of action. Learning or greater understanding is positioned as the interim outcome of inquiry which may inform practitioner action The related concepts of Problem Complexity and MSS Problem Space Match are acknowledged as similarly distinct, recognising the likelihood of multiple Tools and Systems being used to address either or both 'spaces'

Observations	Relevant Related Research	Gap/Deficiency	Contribution/ Claim/ Implication
	Inquiry Logics Research concepts related to inquiry (Blaikie: 2010) Sensemaking and framing (Weick: 1995), cues related to inherited background knowledge (Tsoukas: 2005, 2009), Polanyi: 1966). Cognitive bias (Kahneman: 2011), entrenchment (Arvidsson et al.: 2014) and performance bias (Eisner: 2003) Lack of recognition of Sensing activity (Marchand: 2011) and information overload (Hopkins et al.: 2010, Dunlop: 2009) Importance of Practitioner Alignment as an enabler for learning (Wenger: 1998)	Relevant research inquiry concepts and considerations are not evident in the MSS model, nor the importance of making them explicit in order to keep them under review, or test related assumptions Making this explicit helps avoid entrenched or implicit thinking, and helps clarify and prioritise relevant data, or sensing, which is not currently addressed in the MSS model. It also facilitates alignment and shared understanding of the research framing, objectives, methods, etc. between participants and different practitioner groups involved. However, alignment between practitioner groups involved is not present in the MSS model or highlighted as important	The concept of Data is introduced, as representing a snapshot of what Practitioner Groups deem relevant characteristics about a Phenomenon, in order for them to address prioritised (Research) Questions Validity and measurement concepts in relation to Data are identified as important inquiry considerations An Alignment construct is introduced between different Practitioner Groups involved. **See also Practioner Group observation/contribution**
	Benefit Realisation Logics Benefits realisation (Ward & Daniel: 2012) Theories-in-Use (Argyrus & Schon: 1978), theories-of-action (Weick: 1995) and challenges of cognitive entrenchment (Arvidsson et al.: 2014) and performance bias (Eisner: 2003). Information Orientation (Marchand et al.: 2001) Encouraging ambidexterity (Huang et al.: 2014, Turner, et al.: 2013)	MSS benefits defined abstractly as improved decision-making rather than in contextualised practitioner terms. The importance of making benefit realisation explicit and specific is not addressed Challenges of entrenched practitioner thinking and biases are not included in the MSS model or highlighted as important considerations	Introduction of benefit realisation space concept to distinguish it from the related inquiry being pursued, though recognising an overlap and interdependency (e.g. in engaged inquiry) Different Complexity concepts are introduced related to each 'space' or logic, as different challenges are likely to be encountered. Similarly, different MSS Problem Space Match concepts and indeed different systems are envisaged. Barriers to Tool and Data use are introduced in recognition of related Engagement and practitioner learning challenges

Observations	Relevant Related Research	Gap/Deficiency	Contribution/ Claim/ Implication
Engagement with Data and/or the Phenomenon is central to practitioner learning	Research inquiry (Blaikie: 2010) Action Research (Eden & Huxham: 2002) Design Ethnography (Baskerville & Myers: 2015) Situated Learning (Vera & Crossan: 2003, Wenger: 1998), Sensemaking and theories of action (Weick: 1995, Argyrus & Schon: 1974, 1978) Data availability (Davenport: 2014), Information Orientation (Marchand et al.: 2001) and the variety of bricolage and material approaches (Duymedjian & Ruling: 2010, Stigliani & Ravasi: 2012)	No recognition in MSS of multiple concurrent approaches to learning and action that may be pursued and how they may interact	Direct Engagement with a Phenomenon and Indirect Engagement with related Data are recognised as distinct and potentially complementary approaches in the explanatory model, influenced by practical considerations such as availability and practitioner preferences
The iterative nature of inquiry and situated learning (and related Data refinement)	Research progression from What, to Why, to How questions (Blaikie: 2010), i.e. theory building to theory testing, as well as the provisional nature of knowledge and theory Constant revision and refinement of theories-of-action (Weick: 1995), Absorptive Capacity and path dependency (Cohen & Levinthal: 1990) Increasingly fine distinctions as a potential definition of increasing knowledge (Tsoukas: 2005)	The MSS does not include iterations of refinement or evolution of the MSS, underlying data and related knowledge or theory about a phenomenon. Instead these are effectively treated as constants or givens	Iterations of Engagement recognised in my explanatory model Theory-building versus theory testing and exploitation are distinguished. The former implies longer time-frames and more iterations before theory exploitation and related benefit realisation may be possible Theory Maturity constructs are introduced both in relation to the forms of inquiry and techniques being used, as well as the practitioner knowledge about the Phenomenon and theories of action being pursued

Observations	Relevant Related Research	Gap/Deficiency	Contribution/ Claim/ Implication
Giving Data prominence and disentangling it from wider MSS			
Data and algorithmic entanglement within Tools	Research data analysis supported by distinct tools for particular kinds of data and related analysis (e.g. quantitative and qualitative data and tools) (Blaikie: 2010, Lewins & Silver: 2007) Systems as negotiated reifications of knowledge and economies of meaning (Wenger: 1998), their design and reinterpretation (Orlikowski: 1993), and the socio-materiality of IT artefacts (Cecez-Kecmanovic et al.: 2014, Leonardi: 2013, Orlikowski: 2007, 2006, 2002, 2000, Orlikowski & Iocona: 2001)	Data is not reflected explicitly in the MSS model, nor related availability and engagement. Instead, the level of MSS Tool Use is highlighted as important and not distinguished Ideas of reification and socio-materiality focused research has not focused on data, as distinct from theorising the wider related IT artefact	Algorithmic and Data elements of an MSS are distinguished and Data is introduced as a distinct construct, though recognising Data may be incorporated within a wider MSS. Tools are positioned as mediating access and engagement with embedded data to varying degrees depending on the nature of the Tool and degree to which Data is integrated within it The nature of the Phenomenon is recognised as an important determinant of the kind of Data that can be collected (e.g. qualitative or quantitative) and the forms of Engagement chosen Quality of Data Engagement and related Barriers to Data Engagement are introduced as important *and highlighted as promising areas for further research*

Observations	Relevant Related Research	Gap/Deficiency	Contribution/ Claim/ Implication
Theorising Data – authorship, use and evolution	Data, Information and Knowledge definitions, distinctions and interactions (Kettinger & Li: 2010) and Capta concept (Checkland & Holwell: 1998) Knowing and socially constructed approaches to conceptualising knowledge (Newell & Galliers: 2006, Tsoukas: 1995, Blackler: 1995, 1993), emphasising context bounded nature, implicit precepts and authorship (Tsoukas: 2005) and related category problems (Star: 2010) Importance of sensing activities, data quality and attributes and more related research (Tamm et al.: 2013, Kettinger & Marchand: 2011, Marchand et al.: 2001) Systems as negotiated reifications of knowledge and economies of meaning (Wenger: 1998), design and reinterpretation (Orlikowski: 1993), and socio-materiality of IT artefacts (Cecez-Kecmanovic et al.: 2014, Leonardi: 2013, Orlikowski: 2007, 2006, 2002, 2000, Orlikowski & Iocona, 2001) Boundary artefacts and infrastructures (Star: 2010, Wenger: 1998) and facilitation of emergent knowledge (Tsoukas: 2009)	Ongoing challenges in distinguishing Data, Information and Knowledge constructs, noting relatively interchangeable use of Data and Information terms in the IS literature reviewed Dominant definitions and theorising of Data in IS don't sufficiently address its inherent socially constructed, context bounded nature, and leverage equivalent theorising efforts within Knowlede Management While trust in Data is recognised as an important factor, it is not reflected in the MSS. That it may be fostered through Engagement, which can provide a basis for practioner confidence in claims and as a basis for related action, is also not shown Socio-material researchers have not specifically focused on theorising data, as distinct from wider IT artefacts. A central ongoing debate relates to the fundatmental nature of IT artefacts as equivalent actants and a focus on performative aspects versus a need to recognise a distinction between human and technology actants	Cross case observations richly describe Data's socially constructed nature, authorship and use by different practioner groups involved, often playing the role of a boundary object or infrastructure The related discussion about theorising Data explores the need for a better definition and theory, offering a basis for distinguishing it as physical and persistent compared to embodied Information and Knowledge. This lends support for the importance of distinguishing human and technology actants The discussion explores various promising avenues for further theorising and conceptualising Data on an interdisciplinary basis (especially with Knowledge Management and Learning): • Authorship and evolution of data structure and content, as negotiated economies of meaning • Reinterpretation during use, especially where this is separated from authorship, related practitioner purpose, or temporally (linking this to trust and confidence issues) • Its role as a boundary object or infrastructure, where shared across practitioner groups/functions, and what influences its effectiveness (e.g. barriers) • Accounting for different forms of and degree of mediation of data engagement by tools

Observations	Relevant Related Research	Gap/Deficiency	Contribution/ Claim/ Implication
Addressing different practitioner groups and Knowledge domains			
Different Practitioner Groups pay attention to different cues, contribute and develop different local domain Knowledge	Data Analysts and Managers (Wang & Wang: 2008) Sensemaking (Weick: 1995), Sensing (Kettinger & Marchand: 2011), Cues and Knowledge (Tsoukas: 2005, 2009), Absorptive Capacity-path dependency (Cohen & Levinthal) Communities of Practice, practice boundaries, economies of meaning and reification in systems (Wenger: 1998) and different practitioner domain knowledge (Ribes & Bowker: 2009)	Practitioners are characterised as either tool users/analysts or decision-makers in MSS rather than as a potentially broader set of different practitioner groups working together Other discipline areas are not addressed or reflected in the MSS model, with no mention of practitioner boundaries involved or the particular role of designers	A range of Practitioner Groups is recognised in the explanatory framework, working together and contributing different domain knowledge and skills The knowledge domain related to the central Practitioner Group Phenomenon(a) is positioned as central to generating insights and related benefits. The maturity of existing knowledge and theory related to the Phenomenon is positioned as an important factor for such initiatives The role and institutionalised thinking of those involved in data design and/or related tool selection/ design/ configuration is also highlighted as likely to be important for bounding data and tool mediation
Importance of Alignment and the related role of Boundary Spanning	Communities of Practice, economies of meaning, alignment and boundary spanning (Wenger: 1998) Domain knowledge, shared understanding and related boundary clarification/ networks (Ribes & Bowker: 2009, Edwards et al.: 2011, Cummings & Pletcher: 2011, Balogun et al.: 2005) and related absorptive capacity (Cohen & Levinthal: 1990), facilitated by slack resources (Salge & Vera: 2013) Boundary objects and infrastructures (Star: 2010)	The importance of spanning and alignment of different practitioner groups involved is not recognised in the current MSS model While recognising the importance of boundary engagement processe to generate shared understanding, existing work on shared 'ontologies' focuses on codifying domain metadata rather than recognising its provisional, evolving and socially constructed nature	The number of different practitioner groups involved is recognised as an important dimension of inquiry and benefit space complexity. Related separation is recognised as a potentially important barrier to data engagement Alignment within and across different groups is introduced as an important factor or enabler for emergent learning. The importance of boundary spanners and related activity is emphasised, and making inquiry and benefit logics explicit, thereby available for review, discussion and agreement

Observations	Relevant Related Research	Gap/Deficiency	Contribution/ Claim/ Implication
Adopting a CoP lens to study Data initiatives			
Value of mapping practice communities, boundary spanners, activity and forums and artefacts, for practitioners and researchers	Communities of Practice, economies of meaning, boundary artefacts (Wenger: 1998) and promoting productive dialogue/ alignment (Stigliani & Ravasi: 2012, Tsoukas: 2009) Situated practitioner knowing, their technology use (Orlikowski: 2002, 2000) and its socio-materiality (Cecez-Kecmanovic et al.: 2014, Orlikowski: 2006, 1992), related techniques of zooming in and out (Nicolini: 2009) and network analysis (Cross et al.: 2006) The importance of practitioner settings and characteristics of site-shifting that promote ambidexterity (Huang et al.: 2014)	While the importance of practitioner contexts has been recognised and characteristics of site-shifting are seen as important for promoting ambidexterity, such contexts haven't been defined in terms of practitioner groups and related contexts Most work on knowing and socio-materiality has not focused on data or IT artefacts as negotiated reifications or economies of meaning, or boundary objects/ infrastructure. With the exception of newer ideas related to zooming in and out, most studies have not focused at the group level phenomena and how this interacts within individual level phenomena	Mapping different practitioner communities is illustrated as a promising and pragmatic technique for practitioners and researchers to: • Bound research contexts or help define project scope (in terms of stakeholder groups) • Identify relevant boundary artefacts, related author and user groups and relevant practitioner knowledge domains • Identify boundary spanners, related activity and forums or mechanisms ***The discussion section argues for using and further developing such mapping as a complementary technique to help integrate research focused at individual and group level practices in relation to emergent learning recognising explicit socio-material artefacts and tacit elements involved***

Observations	Relevant Related Research	Gap/Deficiency	Contribution/ Claim/ Implication
Distinguishing practice boundaries, edges and borders	Boundary concept, boundary objects and infrastructures (Star: 2010; Wenger: 1998) Practice boundaries and definition (Hughes et al.: 2007, Wenger: 1998)	The term boundary was originally intended to denote a shared space between practice areas and related objects as organic structures to enable collaboration without consensus, as distinct from boundary infrastructures. However, its subsequent use has drifted and narrowed and distinctions between boundary objects and infrastructures are not clearly addressed in the CoP framework Defining practice areas is recognised as problematic. In addition, while acknowledging that these may not coincide with institutional boundaries and give rise to tensions, these are not addressed in the current CoP framework	The research findings underline the call for further research to address boundaries and related artefacts, contributing two ideas for further development: • Clearer delineation and definition of practice edges, boundary spaces and institutionalised borders (or structural reifications of practice edges) • The quality of permeability or 'hardness' related to engagement across such practice edges or within boundary spaces, and how this relates to the existence and co-evolution of boundary objects and infrastructures, as well as how these facilitate or in the latter case bound or inhibit emergent learning

9 PRACTICAL IMPLICATIONS

Having outlined various theoretical contributions in Chapter 8, this Chapter returns to the business problem outlined in my introductory Chapter. Given the problem of mixed project benefit outcomes for Data initiatives and projects more generally, it explores how the enhanced understanding of exploratory data initiatives gained from my research might be applied to this challenge.

This can be characterised as a governance challenge to ensure appropriate and consistent outcomes from investing resources to pursue insight from Data and related organisational benefits. The governance challenges of Data initiatives are recognised. For example, a fairly recent call for papers (MISQ: 2014) for a special issue on Data analytics challenges, sought contributions addressing various practical management and governance challenges, amongst other topic areas.

Based on the challenges observed by participants in the cases observed, the section outlines various questions that might be posed by project managers of such initiatives, or individuals otherwise accountable for their governance. This typically could include project sponsors or other stakeholders, who might often be called upon to review related business cases for investment or attend project governance forums to review project progress.

Given the exploratory nature and limited scope of my research fieldwork, these practical governance implications and suggestions are necessarily provisional. They aim at aiding practitioners achieve a better understanding of their Data initiatives and the challenges they may face, rather than offering prescriptive advice on how to manage them.

The final section of the chapter provides a tabular summary linking the main field observations highlighted in chapter 7, and related practical challenges noted, provisional implications from my discussion in chapter 8 and the explanatory framework introduced. It also links these back to the governance questions introduced in this chapter, which seek to surface and help practitioners address these.

9.1 Project challenges observed

The starting point for considering possible governance questions to pose, was reflecting on the challenges encountered by the Data initiatives observed, and thinking about what possible governance responses might help address such challenges. These are summarised in Table 9-1.

Table 9-1 Challenges Encountered and Suggested Governance Focus

Challenges Noted	Governance Response
• Framing and prioritising	• Clarify the nature, purpose and scope of data initiatives • Surface the underlying inquiry and value creation logics • Plan iterative cycles of clarifications and test assumptions early to help prioritise & scope
• Encompassing different agendas	• Highlight the critical practice communities involved and insist on early involvement • Identify and manage tensions
• A lack of sufficient focus on data use	• Clarify inquiry and value creation assumptions • Insist on testing them early through practice led prototyping & experimentation
• 'Cross-border' collaborating challenges	• Co-locate exploration and exploitation within practice contexts • Actively support boundary spanners, activity & artefacts
• Tools distract and constrain thinking	• Avoid anchoring in data, systems and existing practice • Don't optimise too early

The rest of this section turns to the explanatory framework presented in Chapter 7, using this to consider and craft appropriate governance questions for Data initiatives, to surface and address the challenges highlighted.

9.2 Some practical Data project governance questions to pose

9.2.1 Do we really understand our project scope and ambition?

For those with governance responsibilities for Data projects, an early challenge is to ensure the project is framed and scoped correctly, in order to ensure budgeting and benefit cases are likely to be realistic. This means establishing a good grasp of the nature of the project being undertaken, which is explored through a series of related subsidiary questions.

How fuzzy is our project logic?

An important determinant of key project scoping variables, e.g. timeframe and resources required, is the level of clarity and certainty about key project parameters. This is characterised by where it lies on a spectrum of 'fuzziness', which is related to the idea of how bounded a problem might be. Two logics need to be considered, although they are often intertwined in practice and any particular project may include either or both and to different degrees. The first is any inquiry logic at work, while the second relates to how new insights might be realised to create value or benefits within a particular practitioner context, for particular stakeholders. Typically there are sequential dependencies between the inquiry and benefit realisation logics.

At one extreme, the initiative might be early exploratory research where the Phenomenon, research questions, relevant Data and logic are still fuzzy and only likely to emerge or be clarified during the project, or are inherently challenging. At the other extreme, the initiative may represent a straightforward, bounded problem where the research design, Data in view, value creation logic and assumptions are all clear. In academic terms, are we building fundamentally new theory or concepts, and collecting new Data, which are all likely to take time, involve several iterations and possible false starts; or are we using or refining existing, well tested theory and concepts, which may enable us to more relatively quickly and confidently progress to generating and acting on insights addressing a clear research Question and problem?

In relation to the benefit realisation logic, is a clear business problem or aspiration in view, with clarity as to the Practitioner Groups involved, and which Groups are likely to benefit? We also need to consider the level of clarity in relation to the Groups likely to be impacted on and required to help facilitate the benefit realisation, and what learning and Knowledge need to be addressed to ensure envisaged benefit outcomes are realised. These might often be via changes and improvements to institutionalised practices and systems.

The greater the level of clarity, the more straightforward the project assumptions as to scope, resource requirements and timeframes are likely to be. In the face of considerable fuzziness and lack of clarity, achieving clarity about these elements could represent interim or staging point project objectives in their own right. These will need to be made explicit and actively monitored.

Which Data dimensions and framework elements are challenging?

For a more detailed grasp of the nature of the fuzziness or challenge inherent in a project, considering the clarity and challenge around specific framework elements introduced in Chapter 7.1 above can be a helpful starting point. For example, how well do we understand the Phenomena in view, the clarity and consensus around practitioner objectives, the questions any inquiry and experimentation is to address, the existence and clarity of required Data, or relevance of existing Data, the familiarity of tools and techniques for analysis that may be appropriate, etc.?

In the case of *big* Data initiatives, it may also be useful to consider Data more carefully and specifically. Three Data dimensions are often identified as differentiated for such initiatives: *volume, variety* and *velocity* (Laney: 2001). It may be useful to consider which dimensions apply to the initiative being considered. These are likely to impact on elements of the framework differently and throw up different challenges and opportunities for Data initiatives. Let us consider each in turn. *Volume* may perhaps be the least problematic, in the sense that it provides mainly a technical processing and storage challenge, with fairly established approaches and solutions available to address these.

Data *variety*, especially in terms of new sources or new Data types, present more challenges and complexity, both in terms of processing and storage, as well as challenges to analyse and integrate them with more structured Data. This might require new techniques and sometimes a blend of qualitative and quantitative approaches. However, this also provides considerable opportunity for triangulation and mixed method research design opportunities (although this requires appropriate validity and epistemological consideration and alignment).

Velocity, especially when combined with volume, is likely to overload human cognition without appropriate visualisation and seems to lend itself to automation. However, this is likely to require applying robust, tested theory and associated models. Even then, continuing practitioner understanding, monitoring, oversight of boundary conditions will be important. In particular practitioners will need to be sensitive to the need to reframe, revisit and retest underlying theory and algorithms, especially for more rapidly shifting social Phenomena (e.g. customer preference). Tackling velocity seems some way down the track from exploratory, theory-building/refinement Data initiatives and more relevant to more exploitative Data initiatives, once theory has been tested and robustness established.

Is it problem focused or does it aspire to creative innovation?

As already highlighted in 8.1.2, projects focused on clearly defined problems are likely to be far more straightforward to address than aspirational initiatives. This is especially so where such initiatives are pursuing innovation or exploring new opportunities that are not particularly clear or bounded.

In the case of exploratory or innovation focused initiatives, governance will need to focus on maintaining an appropriate balance between (initial) divergent thinking on the one hand, to identify and explore options, and on the other hand, (subsequent) efforts to converge on and prioritise the most promising options to explore further. Identifying these trade-offs and phases explicitly within the programme, carefully considering criteria for evaluation options for pursuit, and allowing sufficient time and iterations to diverge and then reach convergence will be important during planning.

Anticipating team dynamics and the inevitable emotional attachment likely to arise in relation to particular options, need careful consideration. This might include whether or not different teams are set to explore promising options with a view to introducing a competitive dynamic. Doing so will require anticipating and planning how to address the inevitable disappointment of 'losing' teams, whose ideas aren't taken forward, to maintain motivation.

How many iterations might we need?

In project terms, the more uncertainty or fuzziness as to any of the components highlighted by the framework, the longer the project may take and the more iterations may be required. Similarly for more innovation centred initiatives. This needs to be explicitly reflected in longer project timeframes, with more staging and review checkpoints, to focus on achieving greater clarity of the research design components, or testing benefit realisation assumptions.

Those with governance responsibilities will need to establish and maintain a clear view of which elements or assumptions are fuzzy or uncertain, remain to be tested or clarified (i.e. how many 'known and unknown, unknowns' are present). This should be explicitly monitored and reflected in project risk assessments and contingencies. Funding models and business cases are also likely to require different approaches, reflecting more options pricing and venture capital portfolio approaches, at least until more clarity is achieved and more traditional cost-benefit analysis is feasible.

Have we identified a sound portfolio of options or avenues to pursue? Are we clear on our (staged) learning objectives?

Portfolio thinking may be useful both within particular Data initiatives and across a number of Data initiatives to balance risk and focus on synergies. This may be particularly true for innovation focused programmes or initiatives, where various options may be pursued, with different levels of likely opportunity, uncertainty and risk characteristics.

Early testing of inquiry and value logics, and related assumptions should be explicitly prioritised in project plans. This is likely to enhance learning (perhaps

more so from failure), as well as avoid costly overinvestment based on untested logic or theory.

Several options, potential research designs, hypotheses should be explored in the face of ambiguity and uncertainty rather than a single, best approach adopted. This avoids 'locking-in' thinking and framing, as well as too much emotional investment and identification with any particular option or approach. Instead, a focus on synergistic learning, across a portfolio of options and benefit opportunities, should be maintained.

9.2.2 Are Practitioner Groups Aligned on the overall objectives?

Several related questions can also be posed to help identify and promote appropriate Boundary Spanning to achieve Alignment:

- Where and how can we co-locate key groups involved?
- Who are our boundary spanners? Are they supported?
- How is targeted spanning activity facilitated and supported?

While adopting a pure research framing, and positioning the activity in an R&D function, may be useful for exploratory inquiry, separation introduces other challenges or trade-offs in relation to practitioner learning and application of the new Knowledge. At InfraDig, it is certainly evident that separating and locating various elements or activities into different specialised Groups, e.g. for Data definition, Data collection and ultimate use, introduces fragmentation of focus, anchors participants in more immediate activities and practices, and displaces the Phenomenon and use from focal attention. This fragmentation of focus or lack of Alignment is often recognised as 'silo behaviour' by participants (see Appendix D.4). This represents a bounded focus on their immediate activity and goals rather than an alignment to wider Data use and benefit realisation through enhanced Knowledge and insight.

Combining inquiry and application is sometimes referred to as ambidexterity (Turner et al.: 2013). Within the strategy-as-practice literature, Huang et al. (2014), argue for co-locating exploration and exploitation activities within practice areas (or 'sites') to increase the likelihood of such ambidexterity

occurring, and in order to prompt ‘site-shifting’ in terms of reframing and improving practices. Where such co-location is not permanently practical, considerable focus is likely to be required to identify natural or existing Boundary Spanners and supporting appropriate Boundary Spanning activities, artefacts and initiatives. Projects and initiatives may also offer a useful if temporary vehicle for co-location.

9.2.3 Have we got practitioner engagement with problems and Data?

Where is the problem or opportunity located? (which practitioners?)

This seems a crucial governance role to derive a workable scope and business case. On the one hand, this helps bound the focus of the initiative, immediately bringing into focus a clear set of Practitioner Groups and the likely benefits that will accrue from the learning and new Knowledge they pursue, realising benefits through improving their practice.

On the other hand, it also highlights other Practitioner or functional Groups that may need to be involved in the inquiry and in realising practice improvements, together with related areas of Boundary Spanning that need to be actively managed and supported. For example Human Resources, IT and Finance may be important to facilitate organisational and community learning, technical support and funding investment respectively.

Prioritisation is also likely to be easier for business practitioners than Data scientists or project managers, unless the latter are close enough to the core Practitioner Group or organisational concerns, competencies and related goals.

Governance should focus on business linkages and related benefit realisation logic, looking for opportunities to facilitate practitioner reflection. Prioritisation should be revisited at successive stages and decision points throughout the initiative to reflect learning. Governance should also encourage and monitor Practitioner Group ownership of and identification with the outcome, which may ultimately be reflected in their readiness to engage with Data, generate insights and identify, adopt and enact related practice improvements.

Is current practitioner domain Knowledge and Data use enough?

It will be important to assess the current Knowledge and sophistication within particular Practitioner Groups involved, in relation to framework elements critical to the inquiry or benefit realisation (e.g. Tool and technique familiarity, domain Knowledge in relation to the Phenomenon in question, current levels of Data use and facility with Data). This will help calibrate how ambitious a learning journey is being pursued for the Practitioner Groups involved.

Another question may be to assess how entrenched and institutionalised practice in particular Practitioner Group areas has become and how this is likely to influence and bound thinking. This accords with work on strategy blindness and cognitive entrenchment (Arvidsson et al.: 2014), hampering broader, second order benefit realisation. One relevant measure here may be the rate of change and innovation evident within different Practitioner Groups and how they compare to similar Practitioner Groups in other organisations or sectors. Identifying and supporting innovators within particular Practice Group areas may be important in this regard, as well as identifying respected practitioners who may be particularly influential within the particular Practitioner Groups involved.

Are engagement and experimentation opportunities adequate?

Sponsors need to ensure opportunities for early and sustained domain Practitioner Group involvement and Data Engagement are sought by initiatives and if necessary created. Boundary opportunities for learning and actively supporting Boundary Spanners, activities and related Boundary Artefacts, between domain areas may be particularly important here, as well as improving general levels of Data familiarity and competencies over time and in parallel. As argued earlier, this is likely to build both trust and confidence in the Data, as well as related ownership of the inquiry and related Data artefacts being created. This needs to identify the Tools and techniques the Practitioner Groups are already familiar with as a starting point.

Time and opportunity for experimentation are also likely to be important where creativity and innovation are sought. This may be challenging in environments

with little ‘slack’ but ensuring appropriate organisational slack is argued as important in the innovation literature (Salge & Vera: 2013).

9.2.4 Are we selecting or optimising Data and Tools too early?

This can apply to both the inquiry or exploration activities, as well as exploitation. As highlighted in the previous section, the starting point for practitioner Data Engagement is likely to be easier with Tools they are familiar with. For most Practitioner Groups, these are more likely to include common end-user and generic Tools, rather than highly specialised Data analytics Tools. The latter often presuppose considerable Tool and technique Knowledge, which requires significant learning investment on the part of Practitioner Groups and which is likely to be seen as a hurdle or Barrier to Data Engagement.

As highlighted earlier, complex Tools and systems are likely to be optimised for particular Purposes and so may bound thinking. To the extent the initiative is pursuing fresh thinking or a new Phenomenon, Tools need to be flexible to facilitate related Knowledge and Data to evolve rapidly and easily. Even though this might not seem *efficient*, it is difficult to optimise solutions until the dimensions and criteria that need to be optimised are much clearer. Otherwise, a solution may be optimised but according to inappropriate criteria or suboptimally for the ultimate use and Purpose converged on in due course. In Data terms, this means the particularly relevant dimensions, attributes and relationships need to have stabilised before significant effort to optimise Tools, Data structures and related algorithmic solutions is attempted.

9.2.5 Are we integrating Data, Knowledge and Learning efforts?

Given the interdependencies highlighted by my findings and discussion above, exploratory Data initiatives warrant careful Alignment of various organisational functions or dimensions, e.g. HR, IT and the functional unit or Practitioner Groups in focus for any practice improvement.

The Practice Groups’ various Data, Knowledge Management and Learning efforts and initiatives all offer additional sources of leverage and opportunities to Align and foster Data Engagement and exploration. This would hopefully lead to

incorporating related learning more fluidly into wider practice improvements in view within the Practitioner Group. Representation of these Groups or functions (e.g. HR) in governance forums and their close involvement in the Data initiative may increase the chances for such Alignment and help ensure that appropriate synergies are identified and realised.

9.3 Summary Practical Implications Table

My research cautions practitioners against uncritically using the existing MSS model to approach such initiatives and systems. It seeks to provide a complementary practitioner centric explanatory framework for such initiatives, highlighting various additional considerations and refinements that may be important. It also points out areas that warrant further research and testing.

While offering an explanatory framework based on two particular cases, rather than aiming at a normative model for all such initiatives, some preliminary and provisional practical implications and related guidance have been outlined. This has been aimed at project managers and sponsors, as a series of governance questions they can pose and explore to better understand their initiatives. It is hoped that this will aid better scoping, planning and execution of such initiatives to enhance project benefits and outcomes realised.

The table below summarises how my research observations link to practitioner concepts/ challenges noted, their practical implications and how the governance questions outlined seek to address these, in conjunction with the explanatory framework provided. In several cases, observations have been grouped together where they result in similar practical challenges and implications.

Table 9-2 Table Summary of Observations, Practical Implications and related Governance Questions

Observations	Related Practical Concepts/ Challenges	Practical Implications	Relevant Governance Questions
Inquiry and benefit logics pursued are not distinguished from each other and are left implicit by case participants The iterative nature of inquiry and situated learning (and related Data refinement)	In practitioner terms this often relates to (project) scoping, objectives and the approach that will be adopted to achieve them, normally reflected in project governance documentation such as project initiation or scope documents and project plans Formalising these elements is useful to align different stakeholder groups and achieve agreement on priorities and approach, also allowing various project assumptions to be challenged and the approach to be refined The following related practical challenges were observed: • Prioritising and scoping initiatives to meet stakeholder expectations • A lack of alignment between different practitioner groups involved • Key assumptions were not challenged, leading to delays and undermining confidence and credibility Pertinent framework/ theory concepts: Inquiry and Benefit realisation spaces, related research design and benefit dependency logics and assumptions, Problem Complexity, Theory maturity, iterations of refinement	Leaving these logics implicit can lead to lack of alignment, delays and loss of credibility and confidence Gaining alignment and a shared understanding seems both more important and more challenging, where data collection, tool design and benefit realisation are split between different practitioner or organisational groups with different priorities Distinguishing inquiry outcomes from related benefit outcomes and dependencies may be useful, especially where different practitioner groups take the lead on these. Research concepts may be particularly useful to apply to inquiry elements (e.g. validity, etc.) More exploratory initiatives addressing new phenomena or building new knowledge are likely to require more iterative refinement and take longer than those applying or refining established knowledge or theory (i.e. exploitation). This impacts on planning and phasing such initiatives. There is value in reassess learning progress, approach, and likely benefits and budget assumptions at different staging points or between iterations. Indeed assumptions should be made explicit and testing them incorporated in plans and stage gates	Do we really understand our project scope and ambition? How fuzzy is our project logic? Which data dimensions and framework elements are challenging? Is it [project] problem focused or does it aspire to creative innovation? How many iterations might we need? Have we identified a sound portfolio of options or avenues to pursue? Are we clear on our (staged) learning objectives?

Observations	Related Practical Concepts/ Challenges	Practical Implications	Relevant Governance Questions
Engagement with Data and/or the Phenomenon is central to practitioner learning	An underlying assumption was noted that data provided would be used, whereas very limited engagement was actually noted and correspondingly little learning occurred. Indeed there was far more focus on data creation and provision rather than on facilitating and encouraging actual practitioner data engagement Practical challenges/issues noted: • Limited practitioner data engagement occurred and limited related learning Pertinent framework/ theory concepts: Data engagement, Barriers to engagement, Tool mediation of engagement	Practitioner engagement with data provided cannot be assumed and needs to be actively encouraged and monitored The explanatory framework highlights a potential practitioner trade-off between direct Engagement with a Phenomenon and Indirect Engagement with related Data, though may represent complementary approaches, influenced by practical considerations such as availability, ease of engagement and preferences Various barriers to data engagement need to be identified and addressed to help facilitate data engagement. Several are identified and more may exist. These may be more significant in complex organisational or project settings Co-creation of data and involvement in data collection may be important to engender a stake in data ownership, confidence in its quality and using it	Have we got practitioner engagement with problems and Data? Is current practitioner domain Knowledge and Data use enough? Are engagement and experimentation opportunities adequate? Are we integrating Data, Knowledge and Learning efforts?

Observations	Related Practical Concepts/ Challenges	Practical Implications	Relevant Governance Questions
Data and algorithmic entanglement within Tools	Projects often start with selecting tools (and/or data). Often tools incorporate a particular data model, bounding what might readily be available to users without significant reworking in end-user tools or access to additional reporting tools Particular tools do not always address the complete data analysis requirement Practical challenges noted: • Tools distract and constrain or anchor thinking • Requests to standardise data and tools • A need for integration or a change in systems • Resistance to adopt new systems or change them • Data ownership tensions, especially where shared or used across practice areas/organisations, or where data collection is separated from use Pertinent framework/ theory concepts: Data engagement, Barriers to engagement, Tool mediation of engagement, Sensing/Scanning, Authorship	A range of tools should be anticipated rather than a single MSS system, taking account of the data in view (e.g. quantative or qualitative), the kinds of analysis required, as well as existing practitioner tool and data use and preferences Particularly for exploratory inquiry initiatives related to a poorly understood or defined phenomenon, care should be taken not to select or optimise tools and data too early, before data requirements are more clearly understood and have stabilised. Tools selected for such initiatives should allow considerable flexibility for a variety of engagement, analysis, as well as data evolution and reorganisation Where tools or systems are used, which include their own data model, this should be carefully evaluated to understand how this may bound or limit inquiry in relation to particular phenomena of interest. This may highlight the need for additional data or tools. Similarly reporting and user interfaces significantly mediate access to data, change authorship processes or accountabilities, this needs to be carefully considered in comparison to existing practitioner preferences and practices to identify any potential barriers this may impose on data engagement	Are we selecting or optimising Data and Tools too early? Which Data dimensions and framework elements are challenging? How fuzzy is our project logic?
Data authorship, use and evolution as important considerations			

Observations	Related Practical Concepts/ Challenges	Practical Implications	Relevant Governance Questions
Theorising Data – authorship, use and evolution	While this emerged from reflection about IS theory rather than practice, several practical challenges noted may well stem from the need for better theoretical ideas about Data: • Importance of scanning activities not recognised • Issues of trust/data quality • Assumption that data provision leads straightforwardly to knowledge • Resistance and practical limits to encoding knowledge Pertinent framework/ theory concepts: Data construct/theory, Sensing/Scanning, Authorship, Data Quality, Distinction between Data and Information-Knowledge	Caution against simply thinking of data as facts. Instead recognise its bounded, provisional and contextual nature, taking care to relate it the phenomena of interest. In particular identify potential dimensions or aspects of the phenomena the data may not address Pay attention to scanning activities to identify relevant Data Carefully consider the authorship and origin data being used, as well as the original purpose for collecting it, phenomena in view and any theory or institutionalised thinking that may have informed what was collected and how it was measurement Take care about adopting absolute concepts of data quality, recognising its often situated or contextual nature, and using broader research validity concepts where these may be more useful Recognise the human, embodied nature of information and knowledge, where managing human framing, sensemaking and related human capital may be particularly important Be careful of simplistic and processing style relationships between data, information and knowledge and related assumptions	Which Data dimensions and framework elements are challenging? Are we selecting or optimising Data and Tools too early? Are we integrating Data, Knowledge and Learning efforts?

Observations	Related Practical Concepts/ Challenges	Practical Implications	Relevant Governance Questions
Different Practitioner Groups pay attention to different cues, contribute and develop different local domain Knowledge Data authorship, use and evolution as important considerations Importance of Alignment and the related role of Boundary Spanning	As highlighted in the cases observed, large data initiatives involve a number of different practitioner groups. This often involves particular groups focusing on data collection, versus system/tool design, versus data use. In project terms these may surface as important stakeholder groups with particular accountabilities Practical challenges noted: • Scoping and prioritisation challenges • Lack of alignment on overall objectives • Lack of clarity as to roles • Lack of ultimate practitioner group engagement • Data ownership tensions, especially where shared or used across practice areas/organisations Pertinent framework/ theory concepts: Practioner Groups' domain Knowledge, Authorship, Theory Maturity, Inquiry & Benefit Space Complexity, Boundary Spanning, Data as a Boundary Artefact, Barriers, reification of economies of meaning, path dependencies for learning and Absorptive Capacity	There is value in clarifying and making explicit which practitioner groups are involved, the particular domain knowledge they will contribute, in relation to overall inquiry, practitioner learning and benefit objectives The maturity of related domain knowledge will impact on the level of effort, timeframes, learning and iterations required for related activities. Actively monitor related learning and development of important areas of domain knowledge periodically, e.g. at project stage gates The practitioner group engaged with the ultimate phenomenon of interest and related benefit realisation should have primacy or take the lead, rather than facilitating practitioner groups, in terms of overall project governance and accountability. This should specifically address questions of shared data ownership and data quality considerations Pay attention to aligning different practitioner groups around ultimate objectives. Identify areas where boundary spanning collaboration and coordination will be important, boundary spanners within related practitioner groups and actively support related activities, forums and boundary spanners	Are Practitioner Groups Aligned on the overall objectives? Where and how can we co-locate key groups involved? Who are our boundary spanners? Are they supported? How is targeted spanning activity facilitated and supported? Have we got practitioner engagement with problems and Data? Where is the problem or opportunity located? (which practitioners?) Is current practitioner domain Knowledge and Data use enough? Are engagement and experimentation opportunities adequate? Are we integrating Data, Knowledge and Learning efforts?

Observations	Related Practical Concepts/ Challenges	Practical Implications	Relevant Governance Questions
Value of mapping practice communities, boundary spanners, activity and forums and artefacts, for practitioners and researchers	As per the previous section	A visual representation of practitioner groups, artefacts, forums, etc. may be easier for practitioners to engage with during planning and related governance activities, especially around stakeholder engagement Various state aseessments and progress indicators can be incorporated in such visualisations in related governance reporting, e.g. heatmaps Improved understanding from further research and theorising boundary spaces, objects and infrastructures, their development and interaction may assist in defining more effective shared data governance, and boundary spanning support mechanisms and interventions	**As per the previous section**
Distinguishing practice boundaries, edges and borders			

10 VALIDITY AND LIMITATIONS

10.1 Validity

The question of how best to demonstrate validity for qualitative or interpretive research is still evolving and under active debate, both within the IS field and more broadly within organisational research (Baskerville & Myers: 2015, Sandberg: 2005, Klein & Myers: 1999).

To demonstrate validity within my research, the following criteria set out by Golden-Biddle and Locke (1993), as cited by Baskerville and Myers (2015) for more traditional ethnographies, have been used as a starting point:

- Authenticity
- Plausibility
- Criticality

Each is considered in terms of how they have been addressed in my research and how I have sought to demonstrate them within the thesis. Through using this structure, I have also sought to touch on new or related thinking on validity where considered relevant and useful.

10.1.1 Authenticity

This concerns the fieldwork undertaken (Schultze: 2000) and providing persuasive evidence of having been immersed in the field and having achieved good access to the phenomenon. I have sought to do so in the following ways:

- Primarily, Chapter 6 provides a rich confessional account of the GoCouncil case and a thick or rich thematic description of the larger InfraDig case.
- Further background and context are provided in Chapter 4 about both cases, together with an overview of my involvement in both. This also addresses what Klein and Myers (1999) term ‘the principle of contextualisation’, to explain how the current situation studied has emerged.

- In addition, Chapters 3 and 5 set out in detail the broad approach adopted, as well the details of the different elements of data analysis performed – some of this co-produced with participants while still immersed in the field.
- Further evidence of fieldwork and data analysis is provided by way of Appendix B and Appendix D.

Though different from action research, Baskerville and Myers (2015) highlight its close affinity to more participative kinds of ethnography. With this affinity in mind, the above approach also aims at addressing what in action research terms (Checkland: 1999) is called *recoverability*. This is described as being somewhere between the strong criteria of *repeatability* and the weak criteria of *plausibility*. Essentially this is to facilitate transparency as to the research process undertaken, so that a reviewer is able to understand how the research was undertaken in a way that allows them to critically evaluate it.

10.1.2 Plausibility

This is fundamentaly concerned with the account presented, both in terms of its contribution to knowledge and in terms of how convincing or believable the account and related arguments are to the academic audience addressed. Though perhaps best judged by the audience addressed, rather than the researcher or author, below I outline how I have sought to address both aspects in turn.

I have sought to clearly outline a clear contribution to theory synthesis, refinement and extension, and to wider research as follows:

- Setting out a clear practitioner problem and related research question in Chapter 1.
- Demonstrating in Chapter 2 how extant research and theory fails to satisfactorily address the practitioner problem and research question, highlighting a clear gap in the research literature.

- Relating my research findings to existing theory and research in my discussion in Chapter 8, to illustrate clear contributions to defined areas of research (see Table 8-1).

I have also actively sought to present early findings and thinking both to fellow-case participants, and the IS research community at conferences and colloquia. This was aimed at testing my interpretation of observations and data, as well as the plausibility of various findings and arguments. This dialectical engagement and feedback was a source for considerable further reflection, often pointing to potentially relevant research to consider, and ultimately helped refine my thinking, argument and presentation. Hopefully this has made for a more plausible and readable account and clearer arguments.

10.1.3 Criticality

Schultze (2000) argues that this aims at a better understanding of the social world or situation examined, as well as the researcher's world, i.e. to better understand others and ourselves.

Criticality is often achieved through researcher reflexivity (Davies: 2008, Singh & Dickson: 2002). This represents a pervasive requirement and was recognised early in my research as important to address explicitly and consistently throughout the research process. Hence, my approach to such reflexiveness was outlined up-front in Chapter 3, when dealing with my research approach. I sought to further evidence and demonstrate reflexiveness within my thesis as follows:

- I have been explicit about my personal background and motivation to pursue this research in the Chapter 1 introduction.
- I have treated my field observations as data during analysis, as well as seeking to be reflexive about the research and data analysis process itself. It was particularly interesting to reflect on how my thinking evolved over time based on observation, analysis and ongoing dialectical engagement with new literature.

- This is reflected in the early realisation of the recursive nature of my research – studying data initiatives pursuing insight, when that was also what I was seeking to do in my research. This culminated in treating my research as a further case for the purpose of the cross-case comparison set out in Chapter 7. This presented particularly rich opportunities for mirroring and contrasting my research process with the data initiatives examined. This resulted in a rich vein of insights as well as a heightened sense of self-awareness about my research.
- The choice of two contrasting cases and their comparison it was hoped would (and indeed did) present opportunities for triangulation across different contexts, as a source of further opportunities for reflexiveness.
- In my writing and presentation I have sought to be transparent first about my involvement and about my sensemaking, including myself as an authorial voice when this was the case. While this perhaps emerges most clearly in the thick case descriptions in Chapter 6, I have sought to do so throughout. This aligns to what Schultze (2000) describes as self-revealing writing and interweaving observational and reported content. Van Maanen (2011) also argues for the importance of a personalised account in Ethnography, while Tsoukas (2005) argues for the importance of narrative knowledge, properly contextualised.

In its pervasiveness, reflexiveness can be challenging to address. I found other criteria, highlighted in relation to broader interpretive studies, helpful as prompts to my reflexiveness (Klein & Myers: 1999, Sandberg: 2005). For example, in plotting my data on a timeline and reflecting on it, in addition to detailed coding in NVivo, I was aware of searching for synthesis (or a coherent whole) and the interdependence of this on the more detailed analysis and findings. Similarly, Sandberg's (2005) concept of transgressive validity was particularly useful in prompting me to consider and be sensitive to the possibility of equivocal interpretations or the unknowable, particularly during data analysis and writing stages. As indicated above, I have sought to signpost such occasions for equivocal interpretation within my accounts, findings and discussion, even though this may come across to the reader as being tentative.

10.2 Limitations

10.2.1 General and inherent

As the cases observed generated relatively little insight leading to significant practice improvements or realising other benefits, the findings and discussion have relatively less to say about how participants generate benefits from insights than how insights are pursued by them.

More broadly, given the nature of the ethnographic, case based approach adopted, my research focused on achieving an explanatory account rather than identifying causal relationships and mechanisms. The findings are clearly grounded in the cases studied and generalizability is not being sought, although the explanations and related theoretical insights may have value in other, similar settings. These are more likely to include other *exploratory* data initiatives than more straightforward data *exploitation* initiatives, where data, theory and their validity in relation to particular phenomena being addressed, are settled and stable.

The study's focus at the project or initiative level, necessarily means that the study findings do not address closely related individual or organisational level aspects of the phenomenon studied (e.g. cognitive aspects). In some areas, possible connections to these have nevertheless been identified during the discussion as valuable avenues for future research and theorising.

While I sought to be reflexive about various aspects of my research, as described in Chapter 3 and in section 10.1 above, the inherent nature of an ethnographic analysis and thesis represents a personal perspective. This personal perspective is informed by my own practical experience as an IT consultant over many years, observation and sensemaking during fieldwork, as well as the theoretical literature engaged with, that together informed my fieldwork, data analysis and writing.

From a theoretical sensitising perspective, Wenger's (1998) CoP Framework and Weick's (1995) sensemaking approach loomed large. Adopting different theoretical lenses would no doubt prompt paying attention to different cues

during observation and fieldwork, as well as during subsequent data analysis and writing. So, while using them proved particularly useful to uncover and explain the social aspects of learning and sensemaking occurring, yielding valuable related insights, their use is nevertheless likely to result in an incomplete view of the phenomenon of interest. Using different theoretical lenses may yet reveal different aspects about the phenomenon studied and might represent useful starting points for thinking about complementary future research.

10.2.2 In relation to specific areas of contribution

Reframing Data Initiatives

While the explanatory framework, based on the cases observed, is presented as a possible starting point for practitioner-centric theorising about (big) data initiatives, it requires further extension and validation across different types of data initiatives.

As already highlighted, the explanatory framework presented has relatively less to say about data use and subsequent value creation, pointing to the need for further empirical research and theorising of this aspect of the framework in particular. For example, this may well identify further barriers to practitioners enacting new knowledge or insights gained and realising related practice or other improvements.

In addition, improvement suggestions to constructs and further inclusions recommended to the MSS model, are based on findings relating to *exploratory* data initiatives, rather than a study of all types of MSS that it seeks to address. This does point to the need for further empirical work on other types of MSS to inform further elaboration and testing of the MSS model, ideally reframed from a practitioner starting point.

For instance, to further illustrate these limitations and the need for further research, Tamm, Seddon and Shanks (2013) identify three pathways to creating value from data using business analytics tools:

- *Advisory Services:* where experience data analytics advisors work with business decision makers to address both *ad hoc* inquiry projects which may be relatively unstructured or unclear, and to address more routine problems. This is consistent with earlier models identified by Wang and Wang (2008).
- *End-user analytics:* this encompasses *ad hoc*, self-directed and relatively routine use of a variety of analytical tools, reports and dashboards to improve evidence based business decision-making, normally based on the provision of an analytics platform.
- *Tool creation:* where these are created and embedded within operational systems to improve analytical capabilities, though still dependent on use.

My thesis argues for a broader view of potential pathways towards value beyond a decision-making focus, especially for exploratory data initiatives, and identifies potential issues and barriers in relation to tool use across all three pathways. However, the use of industrial scale, centralised, automated, algorithmic solutions to generate value did not arise within the case contexts observed, focused as they were on practitioners and emerging practice improvement, and were perhaps closer to the first two pathways they identify.

Such an industrialised, embedded tool creation pathway may be more common for more standardised or routine, and abstracted contexts or problems, such as automated trading in the financial sector, or Internet-based services, where practices and related knowledge lend themselves to algorithmic codification and considerable benefits of speed and scale can be realised from automation. Such a pathway also seems more exploitative than exploratory in nature (my research focus), and may perhaps follow on from more exploratory initiatives. Further empirical research addressing different types of data and analytical tool use, and related value creation pathways, especially in such large-scale

algorithmic data project contexts, would be useful to extend and refine the explanatory framework emerging from this research.

Adopting a CoP lens

The argument presented for the usefulness of adopting the CoP lens for research and practice of data initiatives should not be seen as recommending it as the only or preferred research approach for such initiatives. While it is particularly useful to aid in understanding social aspects of such projects and neatly integrates and addresses artefacts and codification encountered within such contexts, it is complementary to other research approaches. Indeed, the use of different perspectives and lenses is likely to shed light on different aspects of such initiatives.

As already highlighted in the discussion, the CoP framework does not adequately deal with tensions that may arise between practitioner communities and organisational borders or edges (internal and external), especially where such communities span them.

Technical and project organisation aspects of such initiatives are also not a particular focus of the framework and an area where complementary approaches may well be useful, for instance in the area of learning across projects (Easterby-Smith & Lyles: 2003: pp. 29-33).

While the CoP framework usefully considers situated individuals, thereby spanning the individual and community group levels, its primary focus is nevertheless the community or group level. More detailed, individual approaches, such as network analysis or cognitive approaches are likely to be better suited to such a research focus.

Efforts to theorise Data, Information and Knowledge

While an argument is presented for better defining all three basic concepts within IS, and theorising their interrelationship, the findings and insights are most useful in illuminating the Data construct and some of its relationships to Information and Knowledge concepts, rather than just defining Information and Knowledge concepts.

11 CONCLUSION

My research reveals the social 'messiness' of how exploratory data initiatives are actually pursued by participants, in contrast to the neat logics presented as rationales for undertaking them. The explanatory framework that emerges from the reflexive examination of two contrasting cases and my own research effort, provide a useful basis for arguing for several improvements to the current MSS model, which summarises the dominant IS view in relation to such systems:

- Reframing the notion of a Problem Space to recognise different inquiry and benefit realisation domains with their own, though interrelated logics being pursued. These represent theories that need to be refined iteratively as part of the initiative.
- Recognising Data as a central construct within such models, disentangled from algorithmic elements of MSS with Data Engagement rather than tool use as crucial to generating insight.
- Recognising the likely cross-disciplinary and Boundary Spanning nature of such initiatives, with the associated importance of Alignment of the Practitioner Groups involved.

The value of using the CoP framework is illustrated, arguing for some improvements and further research related to Boundaries. Based on the understanding gained from observing my cases, some questions are provided for practitioners in project governance roles to use. These aim to provide them with a better grasp of their projects in order to be able to better manage them and secure envisaged benefits. The research also makes a connection to wider sociomaterial theorising, arguing for an important practical and theoretical distinction between Data on the one hand, which persists independently of actors, and embodied Information and Knowledge on the other. The cases and my own exploratory research also remind us of the following in relation to Data and its use to generate insight: "the map is not the territory" (Korzybski: 1931). There is likely to be an ongoing need for map-makers and map interpreters, especially when approaching an unknown territory or aiming for uncertain destinations.

REFERENCES

Ackoff, R.L. (1989) “From data to wisdom”, *Journal of Applied Systems Analysis*, 16, pp. 3–9

Alavi, M. and Tiwana, A. (2003) “Knowledge Management: The Information Technology Dimension”, In: *Handbook of Organizational Learning and Knowledge Management,* Easterby-Smith, M. and Lyles, M.A. (Editors), Oxford: Blackwell Publishing, pp. 104–121

Argyris, C. and Schön, D.A (1978) *Organizational Learning: A Theory of Action Perspective,* Reading, U.S.A: Addison-Wesley

Argyris, C. and Schön, D.A. (1974) Theory in Practice, San Francisco, U.S.A.: Jossey-Bass

Arnott, D. and Pervan, G. (2008) “Eight key issues for the decision support systems discipline”, *Decision Support Systems*, 44(3), pp. 657–672

Arnott, D. and Pervan, G. (2005) “A critical analysis of decision support systems research”, *Journal of Information Technology*, 20(2), pp. 67–87

Arvidsson, V., Holmström, J. and Lyytinen, K. (2014) “Information systems use as strategy practice: A multi-dimensional view of strategic information system implementation and use”, *Journal of Strategic Information Systems*, 23(1), pp. 45–61

Balogun, J., Gleadle, P., Hailey, V.H. and Willmott, H. (2005) “Managing change across boundaries: Boundary-shaking practices”, *British Journal of Management,* 16(4), pp. 261–278

Baskerville, R.L and Myers, M.D. (2015) “Design Ethnography in information systems”, *Information Systems Journal*, 25, pp. 23–46

BIM Industry Working Group (2011) *BIM Management for value, cost and carbon improvement – Building Information Management (BIM) Working Party Strategy Paper (A report for the Government Client Construction*

Group), available at: http://www.bimtaskgroup.org/wp-content/uploads/2012/03/BIS-BIM-strategy-Report.pdf

Blackler, F. (1995) "Knowledge, Knowledge Work and Organizations: an Overview and Interpretation", *Organization Studies*, 16(6), pp. 1021–1046

Blackler, F. (1993) "Knowledge and the Theory of Organizations: Organizations As Activity Systems and the Reframing of Management", *Journal of Management Studies*, 30(6), pp.863-884

Blaikie, N. (2010) *Designing Social Research: The Logic of Anticipation*, 2nd Edition, Cambridge, United Kingdom: Polity Press

Blaikie, N. (2007) *Approaches to Social Enquiry: Advancing Knowledge*, 2nd Edition, Cambridge, United Kingdom: Polity Press

Blumer, H. (1969) *Symbolic Interactionism: Perspective and Method*, Englewood Cliffs, New Jersey, Prentice-Hall

Boell, S. and Cecez-Kecmanovic, D. (2010) "Attributes of Information", *AMCIS 2010 Proceedings*, Paper 129, available at: http://aisel.aisnet.org/amcis2010/129

Bose, R. (2009) "Advanced Analytics: Opportunities and Challenges", *Industrial Management & Data Systems,* 109(2), pp. 155–172

Cecez-Kecmanovic, D., Galliers, R., Henfridsson, O., Newell, S. and Vidgen, R. (2014) "The Sociomateriality of Information Systems: Current Status, Future Directions", *MIS Quarterly*, 38(3), pp. 809–830

Checkland, P. (1999) Systems Thinking, Systems Practice, John Wiley & Sons, Chichester, England

Checkland, P. and Holwell, S. (1998) *Information, Systems and Information Systems – making sense of the field*, Chichester, England: John Wiley & Sons

Chen, H., Chiang, R.H.L. and Storey, V.C. (2012) “Business Intelligence and Analytics: From Big Data to Big Impact”, *MIS Quarterly*, 36(4), pp. 1165–1188

Chia, R. (2002) “The Production of Management Knowledge: Philosophical Underpinnings of Research Design”, In: *Essential Skills for Management Research*, Partington, D. (Editor), London: Sage, ch.1, pp. 1-19

Clark, T.D., Jones, M.C. and Armstrong, C.P. (2007) “The dynamic structure of management support systems: theory development, research focus, and direction”, *MIS Quarterly*, 31(3), pp. 579–615

Cohen, W.M. and Levinthal, D.A. (1990) “Absorptive Capacity: A New Perspective On Learning And Innovation”, *Administration Science Quarterly*, 35(1), pp. 128–152

Cooper, B., Watson, H. and Wixom, B. (2000) “Data Warehousing Supports Corporate Strategy at First American Corporation”, *MIS Quarterly*, 24(4), pp. 547–567

Cross, R., Parker, A., Prusak, L. and Borgatti, S.P. (2006) “Knowing What We Know: Supporting Knowledge Creation and Sharing in Social Networks”, in *Knowledge Management and Organizational Learning*, Prusak, L. and E. Matson (Editors), Oxford: Oxford University Press

Cummings, J. and Pletcher, C. (2011) “Why Project Networks Beat Project Teams”, *MIT Sloan Management Review*, 52 (3), pp. 75–80

DAMA (2009) *The DAMA Guide to The Data Management Body of Knowledge (DAMA-DMBOK Guide)*, Bradley Beach, U.S.A.: Technics Publications

Davenport, T.H. (2014) *Big data @ work*, Boston, Massachusetts: Harvard Business Press

Davenport, T.H. (2009) “Make Better Decisions”, *Harvard Business Review*, 87(11), pp. 117–123

Davenport, T.H., Harris, J.G. and Morison, R. (2010) *Analytics at Work: Smarter Decisions, Better Results*, Boston, Massachusetts: Harvard Business Press

Davenport, T.H., Harris, J.G., De Long, D.W. and Jacobson, A.L. (2001) "Data to Knowledge to Results", *California Management Review,* 43 (2), pp.117–138

Davies, C.A. (2008) *Reflexive Ethnography: A Guide to Researching Selves and Others*, 2nd Edition, Abingdon, United Kingdom: Routledge

Davis, G.B. and Olson, M.H. (1984) *Management Information Systems: Conceptual Foundations, Structure, and Development,* 2nd Edition, New York, U.S.A.: McGraw-Hill, pp. 199–268

Denzin, N.K. (1971) "The logic of naturalistic inquiry", *Social Forces*, 50, pp. 166-82

D'Eredita, M.A. and Barreto, C. (2006) "How Does Tacit Knowledge Proliferate? An Episode-Based Perspective", *Organization Studies,* 27(12), pp.1821–1841

Dewey, J. (1938) *Logic: The Theory of Inquiry*, New York, U.S.A.: Henry Holt and Company

Dewey, J. (1910) *How We Think*, U.S.A: D.C. Heath & Co

Douglas, J.D. (1971) *Understanding Everyday Life*, London, Routledge and Kegan Paul

Douglas, M.D. and Peppard, J. (2014) "Adopting a situated learning framework for (big) data projects", in *Proceedings of International Data and Information Management Conference (IDIMC 2014)*, Loughborough, United Kingdom, available at http://idimc.org/2014-conference-3/conference-proceedings/

Douglas, M. D. and Peppard, J. (2013) "Theorizing Data, Information and Knowledge constructs and their inter-relationship", In: *UKAIS 2013 conference proceedings*, available at: http://aisel.aisnet.org/ukais2013/

Dunlop, C.A. (2009) “The temporal dimension of knowledge and the limits of policy appraisal: biofuels policy in the UK”, *Policy Sciences*, 43(4), pp. 343–363

Duymedjian, R. and Ruling, C. (2010) “Towards a Foundation of Bricolage in Organization and Management Theory”, *Organization Studies*, 31(2), pp.133–151

Easterby-Smith, M. and Lyles, M.A. (2003) *The Blackwell Handbook of Organizational Learning and Knowledge Management*, Oxford: Blackwell Publishing

Easterby-Smith, M., Thorpe, R. and Jackson, P.R. (2008) *Management Research*, 3rd Edition, London, United Kingdom: Sage Publishing

Eden, C. and Huxham, C. (2002) “Action Research”, In: *Essential Skills for Management Research*, Partington, D. (Ed), London, United Kingdom: Sage Publishing, p. 254

Edwards, P.N., Mayernik, M.S., Batcheller, A.L., Bowker, G.C. and Borgman, C.L. (2011) “Science friction: Data, metadata, and collaboration”, *Social Studies of Scienc*e, 41(5), pp. 667–690

Eisner, A.B. (2003) “The Effects of Rapid Environmental Change on Competitive Strategies: An Organizational Learning Perspective”, *Academy of Strategic Management Journal*, 2, pp. 33–48

Elkjaer, B. (2003) “Social Learning Theory: Learning as Participation in Social Processes”, In: *Handbook of Organizational Learning and Knowledge Management,* Easterby-Smith, M. and Lyles, M.A. (Editors), Oxford: Blackwell Publishing, p. 94

Emerson, R.M., Pretz, R.I. and Shaw, L.L. (2001) “Participant Observation and Fieldnotes”, In: *Handbook of Ethnography*, Atkinson, P., Coffey, A., Delamont, S., Lofland, J. and Lofland, L. (Editors), London: Sage, p. 352

Fekri, H. (2010) *How do Managers Decide Upon the Bases & Methods of Segmentation: A Sensemaking Perspective*, Master's Thesis, Cranfield University

Fontana, A. (1994) "Ethnographic trends in the postmodern era", in Dickens, D.R. and Fontana, A. (Editors) *Postmodernism and Social Enquiry*, London, UCL Press, pp. 203-23

Foss, N.J. and Mahnke, V. (2003) "Knowledge Management: What Can Organizational Economics Contribute?", In: *Handbook of Organizational Learning and Knowledge Management*, Easterby-Smith, M. and Lyles, M.A. (Editors), Oxford: Blackwell Publishing, p. 78

Freire, P. (1970) *Pedagogy of the Oppressed*, New York, Seabury Press

Gadamer, H.G. (1989) *Truth and Method*, 2nd Edition, New York, Crossroad

Geertz, C. (1973) "Thick description: toward an interpretive theory of culture", in Geertz, C. (Editor) *The interpretation of Cultures: Selected Essays*, New York, Basic Books, pp. 1-32.

Gherardi, S. (1999) "Learning as Problem-Driven or Learning in the Face of Mystery?", *Organization Studies,* 20, pp. 101–123

Giddens, A. (1976) *New Rules of Sociological Method*, London, Hutchinson

Giddens, A. (1984) *The constitution of society: Outline of the theory of structure*, Berkeley, California: University of California Press

Golden-Biddle, K. and Locke, K. (1993) "Appealing Work: An Investigation of How Ethnographic Texts Convince", *Organization Science*, 4 (4), pp. 595–616

Guba, E.G. (1978) *Toward a Methodology of Naturalistic Inquiry in Educational Evaluation*, Los Angeles, California, Centre for the Study of Education, UCLA Graduate School of Education

Gummesson, E. (2000) *Qualitative Methods in Management Research*, 2nd Edition, Thousand Oaks, California: Sage Publications Inc.

Habermas, J. (1970) “Knowledge and interest”, in Emmet, D. and MacIntyre, A. (Editors) *Sociological Theory and Philosophical Analysis*, London, Macmillan, pp. 36-54

Hammersley, M. and Atkinson, P. (2007) *Ethnography: Principles in practice*, 3rd Edition, New York, United States of America: Routledge (Taylor & Francis Group)

Harrison, A. (2002) “Case Study Research”, In: *Essential Skills for Management Research*, Partington, D. (Editor), London, United Kingdom: Sage Publishing, pp. 158-180

Hemp, P. (2009) “Death by Information Overload”, *Harvard Business Review*, 87(9), pp.82–90

Hopkins, M.S, Lavalle, S. and Balboni, F. (2010) “The New Intelligent Enterprise 10 Insights: A First Look at the New Intelligent Enterprise Survey on Winning with Data 10 Data Points: Information and Analytics at Work”, *MIT Sloan Management Review*, 52(1), pp. 22–31

Huang, J., Newell, S., Huang, J. and Pan, S. (2014) “Site-shifting as the source of ambidexterity: Empirical insights from the field of ticketing”, *Journal of Strategic Information Systems*, 23, pp. 29–44

Hughes, J., Jewson, N. and Unwin. L. (2007) *Communities of Practice – Critical Perspectives*, Abingdon, UK: Routledge

ISACA (2013) *COBIT 5: Enabling Information*, ISACA, Rolling Meadows, USA, available at: http://www.isaca.org/COBIT/Pages/COBIT-5-Enabling-Information-product-page.aspx (January 2016)

Jourdan, Z., Rainer, R.K. and Marshall, T. (2008) “Business Intelligence: An Analysis of the Literature”, *Information Systems Management*, 25(2), pp. 121–131

Kahneman, D. (2011) *Thinking fast and slow*, London: Allen Lane (Penguin)

Kettinger, W.J. and Li, Y. (2010) “The Infological Equation Extended: Towards Conceptual Clarity in the Relationship between Data, Information and Knowledge”, *European Journal of Information Systems*, 19 (March), pp. 409–421

Kettinger, W.J. and Marchand, D.A. (2011) “Information Management Practices (IMP) from the Senior Manager’s Perspective: an Investigation of the IMP Construct and its Measurement”, *Information Systems Journal*, 21, pp. 385–406

Klein, H.K. and Myers, M.D. (1999) “A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems”, *MIS Quarterly*, 23(1), pp. 67–94

Korzybski, A. (1931) “A Non-Aristotelian System and its necessity for rigour in Mathematics and Physics (presented at meeting of the AAAS)”, Published as *Supplement III, Science & Sanity*, pp. 747–761

Laney, D. (2001) Application Delivery Strategies: 3D Data Management – Controlling Data Volume, Velocity and Variety. Stamford, U.S.A: Meta Group. Available at: http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf

Lave, J. and Wenger, E. (1991) *Situated Learning: Legitimate Peripheral Participation*, Cambridge: Cambridge University Press

Lee, A.S., Thomas, M. and Baskerville, R.L. (2015) “Going back to basics in design science: from the information technology artifact to the information systems artifact”, *Information Systems Journal*, 25, pp. 5–21

Leonard D. and Swap, W. (2006) “Generating Creative Options”, in *Knowledge Management and Organizational Learning*, Prusak, L. and Matson, E. (Editors), Oxford: Oxford University Press

Leonardi, P.M. (2013) “Theoretical foundations for the study of sociomateriality”, *Information and Organization*, 23(2), pp. 59–76

Lewis, A. and Silver, C. (2007) *Using Software in Qualitative Research*, London: Sage Publications

Lindblom, C. (1959) “The science of “muddling through””, *Public Administration Review*, 19(2), pp. 79–88

Lofland, J. (1967) “Notes on naturalism”, *Kansas Journal of Sociology*, 3, pp. 45-61

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Byers, A.H. (2011) *Big data: The next frontier for innovation, competition, and productivity*, McKinsey Global Institute Reports, available at: http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation (last accessed November 2015)

March, J.G. (1991) “Exploration and Exploitation in Organizational Learning”, *Organization Science*, 2(1), pp.71–87

March, J.G. (1987) “Ambiguity and Accounting: The Elusive Link between Information and Decision-making”, *Accounting, Organization and Society*, 12(2), pp. 153–168

Marchand, D.A., Kettinger, W.J. and Rollins, J.D. (2001) *Information Orientation – The Link to Business Performance*, Oxford: Oxford University Press

Marchand, D.A. and Peppard, J. (2013) “Why IT Fumbles Analytics”, *Harvard Business Review,* 91(1), pp. 104–112

Matza, D. (1969) *Becoming Deviant*, Englewood Cliffs, New Jersey, Prentice-Hall

Mies, M. (1983) “Towards a methodology for feminist research”, in Bowles, G. and Klein, R.D. (Editors) *Theories of Women’s Studies*, London, Routledge and Kegan Paul, pp. 117-39

Miles, M.B. and A.M. Huberman (1994) *Qualitative Data Analysis*, 2nd Edition, California, U.S.A.: Sage Publishing

Mintzberg, H. (1975) The Manager's Job: Folklore and Fact, *Harvard Business Review*, 53(4), p. 49

MISQ (2014) "Call for Papers – MISQ Special Issue on Transformational Issues of Big Data and Analytics in Networked Business", Editors: Baesens, B., Bapna, R., Marsden, J.R., Vanthienen, J. and Zhao, J.L., *MIS Quarterly*

Newell, S., and Galliers, R. D. (2006) "Facilitating – or inhibiting – knowing in practice", *European Journal of Information Systems*, 15, pp. 441–445

Nicolini, D. (2009) "Zooming In and Out: Studying Practices by Switching Theoretical Lenses and Trailing Connections", *Organization Studie*s, 30(12). pp. 1391–1418

Nonaka, I. (1994) "A Dynamic Theory of Organizational Knowledge Creation", *Organization Science*, 5(1), pp. 14–37

Nonaka, I. Toyama, R. and Konno, N. (2000) "SECI, Ba and Leadership: a Unified Model of Dynamic Knowledge Creation", *Long Range Planning*, 33(1), pp. 5–34

Orlikowski, W.J. (2007) "Sociomaterial Practices: Exploring Technology at Work", *Organization Studies*, 28(9), pp. 1435–1448

Orlikowski, W.J. (2006) "Material Knowing: the Scaffolding of Human Knowledgeability", *European Journal of Information Systems*, 15(5), pp. 460–466

Orlikowski, W.J. (2002) "Knowing in Practice: Enacting a Collective Capability in Distributed Organizing", *Organization Science*, 13(3), pp. 249–273

Orlikowski, W.J. (2000) "Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations", *Organization Science,* 11(4), pp. 404–428

Orlikowski, W.J. (1992) "The Duality of Technology: Rethinking the Concept of Technology in Organizations", *Organization Science*, 3(3), pp. 398–427

Orlikowski, W.J. and Iacono, C.S. (2001) “Research Commentary: Desperately Seeking the ‘IT’ in IT Research – A Call to Theorizing the IT Artefact”, *Information Systems Research*, 12(2), pp. 121–134

Orlikowski, W.J. and Scott, S.V (2008) “Sociomateriality: Challenging the Separation of Technology, Work and Organization”, *The Academy of Management Annals*, 2(1), Routledge, pp. 433–474

Osterwalder, A. and Pigneur, Y. (2010) *Business Model Generation*, New Jersey, U.S.A.: John Wiley and Sons

Outhwaite, W. (1975) *Understanding Social Life: The Method Called Verstehen*, London: Allen and Unwin

Polanyi, M. (1966) *The Tacit Dimension*, Chicago, The University of Chicago Press

Ranjan, J. and Bhatnagar, V. (2011) “Role of Knowledge Management and Analytical CRM in Business: Data Mining based Framework”, *The Learning Organization*, 18(2), pp. 131–148

Ribes, D. and Bowker, G.C. (2009) “Between meaning and machine: Learning to represent the knowledge of communities”, *Information and Organization*, 19(4), pp. 199–217

Rowley, J. (2007) “The wisdom hierarchy: representations of the DIKW hierarchy”, *Journal of Information Science,* 33(2), pp. 163–180

Salge, T.O. and Vera, A. (2013) “Small Steps that Matter: Incremental Learning, Slack Resources and Organizational Performance”, *British Journal of Management*, 24(2), pp. 156–173

Sandberg, J. (2005) “How do We Justify Knowledge Produced Within Interpretive Approaches”, *Organizational Research Methods*, 8(1), pp. 41–68

Schön, D.A. (1991/1983) *The Reflective Practitioner: How Professionals Think in Action*, London, United Kingdom: Ashgate Publishing Limited

Schultze, U. (2000) “A Confessional Account of an Ethnography About Knowledge Work”, *MIS Quarterly*, 24(1), pp. 3–41

Seely Brown, J. and Duguid, P. (2000) *The Social Life of Information*, Boston: Harvard Business School Press

Shollo, A. and Galliers, R. (2013) "Towards An Understanding Of The Role Of Business Intelligence Systems In Organizational Knowing". *ECIS 2013 Completed Research. Paper 164*.
URL: http://aisel.aisnet.org/ecis2013_cr/164

Singh, V. and Dickson, J. (2002) “Ethnographic Approaches to the Study of Organizations”, In: *Essential Skills for Management Research*, Partington, D. (Ed.), London, United Kingdom: Sage Publishing, p. 117

Snowden, D.J. and Boone, M.E. (2007) “A Leader’s Framework for Decision Making - Wise executives tailor their approach to fit the complexity of the circumstances they face”, *Harvard Business Review,* November 2007, pp. 68–76

Stamper, R. (1991) “The Semiotic Framework for Information Systems Research”, in: *Information Systems research: Contemporary Approaches & Emergent Traditions*, Nissen, H-E., Klein, H.K. and Hirschhaim, R. (Editors), Amsterdam: North Holland, pp. 515–528

Star, S.L. (2010) “This is Not a Boundary Object: Reflections on the Origin of a Concept”, *Science, Technology & Human Values*, 35(5), pp.601–617

Stigliani, I. and Ravasi, D. (2012) “Organizing thoughts and connecting brains: Material practices and the transition from individual to group-level prospective sensemaking”, *Academy of Management Journal*, 55(5), pp. 1232–1259.

Tamm, T., Seddon, P. and Shanks, G. (2013) “Pathways to Value from Business Analytics”, In: *ICIS 2013 Proceedings*. URL: http://aisel.aisnet.org/icis2013/proceedings/KnowledgeManagement/14/

Thorp, J. (2003) *The Information Paradox – Realizing the Business Benefits of Information Technology*, Revised Edition, Mc Graw-Hill Ryerson

Tsoukas, H. (2009) “A Dialogical Approach to the Creation of New Knowledge in Organizations”, *Organization Science*, 20(6), pp. 941–957

Tsoukas, H. (2005) *Complex Knowledge: Studies in Organizational Epistemology*, Oxford, United Kingdom: Oxford University Press

Tsoukas, H. and Vladimirou, E. (2001) “What is organizational knowledge?”, *Journal of management studies,* 38(7), pp. 973-993

Turner, N., Swart, J. and Maylor, H. (2013) “Mechanisms for Managing Ambidexterity: A Review and Research Agenda”, *International Journal of Management Reviews*, 15, pp. 317–332

Van Maanen, J. (2011) *Tales of the Field: On Writing Ethnography*, 2nd Edition, Chicago, United States of America: University of Chicago Press

Vera, D. and Crossan, M. (2003) “Organizational Learning and Knowledge Management: Toward an Integrative Framework”, In: *Handbook of Organizational Learning and Knowledge Management*, Easterby-Smith, M. and Lyles, M.A. (Editors), Malden: Blackwell Publishing, p. 122

Wang, H. and Wang, S. (2008) “A Knowledge Management Approach to Data Mining Process for Business Intelligence”, *Industrial Management & Data Systems*, 108(5), pp. 622–634

Ward, J. and Daniel, E. (2012) *Benefits Management: How to Increase the Business Value of Your IT Projects*, John Wiley & Sons

Weber, M. (1958) *The Theory of Social and Economic Organization*, trans. Henderson, A.M. and Parsons, T., New York: Free Press

Weick, K. (1995) *Sensemaking in Organizations,* Thousand Oaks, California: Sage Publications

Wenger, E. (1998) *Communities of Practice: Learning, Meaning & Identity*, New York: Cambridge University Press

Wixom, B.H. and Watson, H.J. (2001) “An Empirical Investigation of the Factors Affecting Data Warehousing Success”, *MIS Quarterly*, 25(1), pp. 17–41

Yeoh, W. and Koronios, A. (2010) “Critical Success Factors for Business Intelligence Systems”, *Journal of Computer Information Systems*, 50(3), pp. 23–32

APPENDICES

Appendix A Research Design Overview

A.1 Research Design, reflecting final Research Question

	Research Question	Nature of Question	Logic of Inquiry/Research Paradigm		Theory/Concepts	Detailed Design Data Source(s)	Data collection	Proposed Analysis
		(Blaikie: 2010)	Epistemology	Ontology				
1	How do participants in data initiatives actually pursue insight?	What Why	Abductive (Blaikie: 2007) Social Constructionism (Easterby-Smith, Thorpe & Jackson: 2002) Ethnomethedology (Davies: 2008, Sing & Dickson: 2002, Van Maanen: 2011)	Idealist (Blaikie: 2007) Nominalism (Easterby-Smith, Thorpe & Jackson (2002)	Wenger (1998) situated learning framework (Team/group level) Weick (1995) sensemaking (participant level) Blaikie (2007) question progression from what to why to how	Researcher observation Participants Project Records	Diary/journal notes Meeting transcripts/summaries/ recordings Interview transcripts/summaries/ recordings Project Documentation Participant validation	Coding of data using Theory/Concepts identified as a priori coding structure Various forms of Interpretive analysis, relating back to theory & participant feedback: - Question analysis (PoshCouncil) - CoP mapping (visual) (InfraDig) - Narrative Case Analysis (timeline) - Initial Comparative Analysis - Reflexive cross-case analysis
					Various other theories may also be drawn on depending on what emerges during analysis: - material entanglement/tools (Orlikowski) - conversation analysis (Tsoukas) - boundary docs (Tsoukas, Stigliani&Ravasi) - path dependency (Cohen & Levinthal) - Dewey, Elkmaer (directed/exploratory inquiry)			

A.2 Earlier Research Questions (reflecting evolving focus)

What is revealed by using a CoP lens for Data Initiatives as important for enhancing insight and related benefits?

(3rd Review stage, initial drafting of thesis findings-conference papers – June 2015)

How do teams 'frame' data initiatives to generate insight from data?

(2nd Review stage, preliminary data analysis – January 2014)

How do [customer*] analytics project teams use questions to draw insights from [customer*] data?

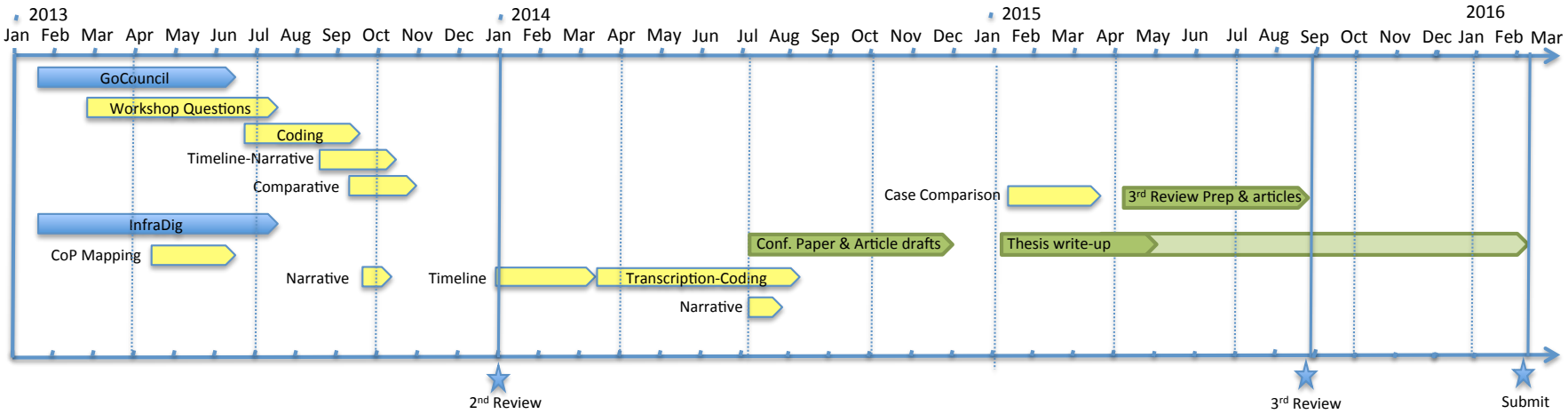
(1st Review stage, based on initial literature review – October 2012)

* At this stage I was considering limiting the research scope to a customer analytics context to make the scope more manageable

Are the theory constructs and hypotheses put forward by Marchand et al. (2001) supported by the evidence of experiences in the context of CRM, BI and KM systems? Which additional constructs and hypotheses are important to include in a more comprehensive theory that seeks to address some of the shortcomings highlighted in the problem description section?

(Research Proposal stage, based on limited reading – May 2010)

A.3 Detailed Timeline of Data Analysis undertaken and Reflexive write-up



Appendix B Interview Protocol

Interview Question	Data to be collected	Purpose	Phenomenon Aspect Explored	Literature Basis
Introduction… A preamble, setting the scence for the interview, introduce myself, provide some background on the research being undertaken and obtaining their permission to a) participate in the research, b) the interview, and in particular c) agreement to record the interview. Assurances as to anonymity and that the research findings will be validated with participants in due course	Interview identification, place and timing Acknowledgement as to understanding Agreement to participate Agreement for the interview to be recorded	Identify recording for analysis Meet Research Ethical Guidelines	**N/A**	
By way of background and introduction…				
Can you tell me a bit about yourself by providing a brief outline your background and experience a) with the company b) similar initiatives/projects c) broader education, training and professional involvement	Company Experience Data Analytics Experience Knowledge (esp. about Data Anal., Tools, phenomenon)	Put them at ease/start with a straightforward question Capture their prior knowledge and experience	**Prior Knowledge**	Weick (1995) - Sensemaking Cohen & Levinthal (1990) - Path dependency (R&D) Wenger (1998) - Situated learning (practice community) Tsoukas (2005) - Inherited background
What is your understanding of the Initiative underway, in terms of its scope and overall objective(s)	Scope understood Objectives understood	Capture their understanding of the purpose(s) of the initiative Understand how they are framing its scope & purpose Identify and understand differences across participants	**Framing of Project Purpose**	Blackler (1995) - Vygotsky's communal (social) model of knowledge creation Wenger (1998) - Power within Practice Communities
Why do you believe it is important? (For the organisation and personally)	Organisational drivers Personal motivations & interests Tensions and politics (personal & organisational)	Identify any personal motivations Start to identify any possible politics and tensions at play	**Participant Purpose(s)** **Power/Social**	
Turning to what you are trying to find out…				
What phenomenon are you trying to find out more about, which is at the heart of the initiative?	Phenomenon understanding	Undestand their framing of the phenomenon of interest	**Framing of Phenomenon**	Weick (1995) - Sensemaking Blaikie (2007 & 2010) - Research phenomenon & Research Questions
What are the dimensions or aspects about this you are most interested in understanding better? OR what ideas about it do you want to test or confirm?	Key dimensions of the phenomenon Hypotheses/models being tested			
How would you frame these as questions you are seeking answers to?	Research Questions	Identify and capture the questions being posed	**Questions**	
Turning to likely logistics….				
What types of Data do you think is likely to be: a) relevant (and why you think so) b) Are they readily available? If so, from where, otherwise where do you think they can be sourced? c) How easy, difficult, expensive is data collection likely to be? Why? How do you anticipate these Data will be analysed (e.g. techniques, approaches) How familiar are you with these approaches? Have you used them before? Describe when and to what extent	Data Types +Participant reasons for relevance Data Sources envisaged Techniques envisaged +Experience with techniques envisaged	Identify and capture the data, techniques and tools likely to be used Understand any starting biases participants bring (e.g. epistemological/epistemological biases and if this is based on prior knowledge or experience versus pragamatic considerations) Are knowledge/experience gaps identified by participants and identify how they plan to address these	**Data Types & Sources** **Techniques/Tools** **Prior Experience influence** **Participant knowledge gaps**	Blaikie (2007 & 2010) - Research Data & Techniques Orlikowski (2006, 2007) - Socio-materiality of tool-use & related knowledge Weick (1995) - Sensemaking Cohen & Levinthal (1990) - Path dependency (R&D) Wenger (1998) - Situated learning (practice experience) Tsoukas (2005) - Inherited background
Is the team likely use any tools to assist in the data analysis? If so, which ones and why them? Have you or the team used these tools before? Describe when and to what extent. Will any training be required/available/provided?	Tools envisaged +Familiarity/experience with tools envisaged Training requirements/knowledge gaps			

Interview Question	Data to be collected	Purpose	Phenomenon Aspect Explored	
Coming back to you…				
Why you believe you've been selected for this project/initiative?	Reason for selection	Identify if organisational and/or domain knowlede has been used as a criteria	**Specialist knowledge**	Sensemaking (1995) - Prior experience, Frames and meaning enactment Wenger (1998) - Practitioner Experience, Group Standing, Practice group structure-situation Tsoukas (2005) - Inherited background, Framing
What will your role be on the project/initiative?	Role within the team	Begin to get an understanding of the team structure, context and likely role understandings and likely dynamics	**Particpant Project Role** **Team structure**	Wenger (1998) - Practice group structure-situation, Likely to
What aspects of your prior experience and knowledge do you think will be useful on this project and why?	Useful prior knowledge/experience +Reason for usefulness of knowledge/experience +How it will be used +When it will be used	Identify to what extent prior knowledge and experience has been used as a criteria for team selection AND to what extent participants are aware of relevant knowledge & experience and how and when this is likely to be useful	**Prior Knowledge**	Weick (1995) - Sensemaking, prior experience framing Cohen & Levinthal (1990) - Path dependency (R&D) Tsoukas (2005) - Inherited background framing
When and how do you think you will bring this prior experience and knowledge to bear on the project	(self-identified)			
Conclude…				
Thank you for your time and the information you've provided. I look forward to working with you on the XXXX initiative over the next few months.	Rapport/response (positive, negative, indifferent)	Start building a rapport to set the stage for ongoing ethnographic involvement	(Likely) Team/Social Dynamic	

Appendix C Data Collected

C.1 Summary of Data Collected

	GoCouncil	InfraDig	Total
Days on Site (Diary)	**17 days**	**31 days**	**48 days**
Participant Interviews	**4 (4:29)**	**10 (9:22)**	**14 (13:51)**
Recorded	3 (3:14)	9 (8:22)	12 (11:36)
Project Meetings *	**9 (10:54)**	**25 (26:28)**	**34 (37:22)**
Recorded	8 (9:54)	11 (13:13)	19 (23:07)
Other Meetings *	**8 (9:53)**	**15 (16:43)**	**23 (26:36)**
Recorded	6 (8:23)	3 (2:18)	9 (10:41)
Total Meetings (no./hrs)	21 (25:16)	50 (52:33)	71 (77:49)
Total Recorded	17 (21:31)	23 (23:55)	40 (45:26)
Total Transcribed/Summarised	12 (14:05)	21 (22:19)	33 (36:24)
Workshops	**2 *(6:00)***	**1 *(2:00)***	**3 *(8:00)***
Artefacts	**27**	**41**	**68**
- Collected	16	33	49
- (Co-) Produced	11	8	19

* Excludes informal collaboration and meetings where no notes were taken – mainly Andrea (GoCouncil)/Barbara (InfraDig)

C.2 Interview Details

	Date	Person	Role	Location	Duration	Record.	Transcr./ Summ.
GoCouncil							
1	08.02.13	Albert	New Business Model Project Manager	GoCouncil, Quiet foyer	*1:15*	-†	Notes
2	08.02.13	Tanya	Performance Reporting Project Team Member	GoCouncil, Meeting Room	1:40	✔	✔
3	19.04.13	Andrea	Johneting Insight Lead	GoCouncil, Meeting Room	0:34	✔	✔
4	26.04.13	Helga	Johneting Facilitator - Planning	GoCouncil, Meeting Room	1:00	✔	✔
Subtotal					4:29	**3/4**	**3/3**
InfraDig							
1	11.02.13		Lead Business Analyst (IS)	InfraDig, Breakout area (booth)	0:44	✔	✔
2	11.02.13		Project Engineer-Contractor	InfraDig, Breakout area (booth)	0:42	✔	✔
3	14.02.13	Mallory	Chief Engineer	InfraDig, Breakout area (booth)	1:02	✔	✔
4	14.02.13	Anthony	IT Systems Development Manager	InfraDig, Breakout area (booth)	1:06	✔	✔
5	14.02.13	Ralph	Asset Data Team Lead	InfraDig, Breakout area (booth)	*1:00*	-†	Notes
6	18.02.13	John	Asset Data Team member	InfraDig, Breakout area (booth)	0:44	✔	✔
7	21.02.13	Paul	KPI Project Manager (Consultant)	InfraDig, Breakout area (booth)	0:32	✔	✔
8	21.02.13		IT Architect/DBA	InfraDig, Breakout area (booth)	0:52	✔	✔
9	28.02.13	James	Asset Data Consultant	InfraDig, Breakout area (booth)	1:12	✔	✔
10	28.02.13	Matthew	Asset Data Super-User	InfraDig, Breakout area (booth)	1:28	✔	✔
Subtotal					9:22	**9/10**	**9/9**
Totals					13:51	**12/14**	**12/12**

KEY:

✔ Transcribed or summarised
Notes – taken at the meeting
† Problem with recording
Italics – estimated duration

C.3 Other Meetings and Workshop Details

	Date	Meeting	Attendees	Location	Duration (hrs:mins)	Recorded	Transcribed/ Summarised
GoCouncil							
MI Project Meetings							
1	15.01.13	Introductory Meeting	Joe (supervisor), Andrea, Patricia	Cranfield University, Joe's office	*1:00*	-	Notes
2	22.01.13	Kick-off Meeting	Albert, Andrea, Patricia	GoCouncil, Patricia's office	*1:30*	✔†	✔
3	08.02.13	MI Review Meeting	Albert, Andrea, Patricia, Tanya, Alex (marketing team)	GoCouncil, Patricia's office	1:32	✔	✔
4	12.04.13	MI Review Meeting	Patricia, Andrea, Albert, Tanya & Alex	GoCouncil, Patricia's office	1:14	✔	✔
5	12.04.13	Engine Room Briefing	Patricia, Andrea, Albert	GoCouncil, Patricia's office	0:43	✔	✔
6	19.04.13	MI Review Meeting	Patricia, Andrea & Alex	GoCouncil, Patricia's office	1:34	✔	✔
7	26.04.13	Engine Room Discussion	Justin (Dep.CEO), Alex, Albert, Andrea, Tanya	GoCouncil, Patricia's office	1:02	✔	✔
8	10.05.13	MI – Building Control (BC) Kickoff	Albert, Helga, BC Leadership team	GoCouncil, Patricia's office	1:13	✔	✔
9	10.05.13	MI – Community Centres (CC) Kickoff	Alex, Albert, CC lead	GoCouncil, Engine Room	1:06	✔	✔
Subtotal					**10:54**	**8/9**	**8/8**
Other Meetings							
1	08.03.13	IT involvement	Dennis (CIO)	GoCouncil, Alan's office	0:57	✔	✔
2	15.03.13	CEO Update	Alexander (CEO), Andrea	GoCouncil, Alexander's office	2:21	✔	Notes
3	26.04.13	CRM Requirement	Complaints lead	GoCouncil, Meeting Room	0:21	✔	Notes
4	03.05.13	IT involvement (follow-up)	Dennis (CIO) & Patricia	GoCouncil, Alan's office	0:54	✔	Notes
5	29.05.13	Enough MI	Patricia, Internal Audit/Risk	GoCouncil, Patricia's office	*1:00*	-	Notes
6	05.06.13	CRM Meeting	Patricia, Dennis (CIO), various IS colleagues	GoCouncil, Engine Room	1:35	✔	Notes
7	27.09.13	Informal Catch-up	Patricia, Alex	GoCouncil, Cafe Area	*0:30*	-	-
OtherCouncil							
8	04.06.13	**OtherCouncil Customer Analytics**	Head of Customer Analytics	OtherCouncil, desk	2:15	✔	Notes
Subtotal					**9:53**	**6/8**	**1/6**
Johnet Insight (MI) Workshops							
1	22.02.13	**MI Workshop (Leisure)**		GoCouncil, Conference Room	*3:00*	-	-
2	14.06.13	**MI Workshop (Community Centres)**	Alex	GoCouncil, Engine Room – Adjacent Meeting Room	*3:00*	-	-
Totals					**26.47**	**14/17**	**9/14**

	Date	Meeting	Attendees	Location	Duration (hrs:mins)	Recorded	Transcribed/ Summarised
InfraDig							
Project Meetings (Reporting & Strategy initiatives)							
1	24.01.13	Big4 Info Strategy catch-up	Donald, 2 consultants	InfraDig, Meeting Room	*0:45*	-	Notes
2	11.02.13	Data MI Project (Kick-off)	Ralph, Anthony, John, Matthew, James, Paul	InfraDig, Meeting Room	1:28	✔	✔
3	14.02.13	Data MI Project Meeting 2 (Design)	Ralph, John, Matthew, James, Paul, Mallory	InfraDig, Meeting Room	1:10	✔	✔
4	14.02.13	Info Strategy Catch-up	Donald	InfraDig, Breakout area (booth)	0:52	✔	✔
5	18.02.13	Data MI Project Meeting 3 (Review)	Ralph, Anthony, John, Matthew, James, Paul, Mallory	InfraDig, Meeting Room	1:28	✔	✔
6	26.02.13	Data MI Project Meeting (Final)	Ralph, Anthony, John, Matthew, James, Paul, Mallory	InfraDig, Meeting Room	0:55	✔	✔
7	12.03.13	Info Strategy (brief catch-up & CoP intro)	Mallory (Donald stuck in the snow)	InfraDig, near Mallory's desk	*0:45*	-	Notes
8	14.03.13	Info Strategy (catch-up & CoP intro)	Donald	InfraDig, Breakout area (booth)	*1:00*	-	Notes
9	28.03.13	Info Strategy Meeting (catch-up)	Donald	InfraDig, Breakout area (booth)	0:42	✔	✔
10	09.04.13	Document Controllers workshop	Barbara, Document Control Manager, Matthew, various others (30+)	InfraDig Training Room	2:00	-	Notes
11	09.04.13	Info Strategy (catch-up and intro)	Barbara, Document Control Manager	InfraDig Breakout area (booth)	*0:30*	-	Notes
12	18.04.13	Info Strategy (CoP-Asset view)	John & Matthew	InfraDig Breakout area (booth)	*1:00*	-	Notes
13	18.04.13	Info Strategy (catch-up)	Barbara, Mallory	InfraDig Breakout area (booth)	*0:30*	-	Notes
14	22.04.13	Info Strategy (CoP-Broader view)	James (Asset Consultant)	InfraDig Breakout area (booth)	*1:00*	-	Notes
15	25.04.13	Info Strategy (CoP-Asset follow-up)	John (Asset Data team)	InfraDig Breakout area (booth)	*0:45*	-	Notes
16	25.04.13	Info Strategy (Prezi CoP review)	Barbara	InfraDig Breakout area (booth)	*1:00*	-	Notes
17	01.05.13	Info Strategy (CoP IS view)	Anthony & IT Business Analyst	InfraDig Breakout area (booth)	*1:00*	-	Notes

	Date	Meeting	Attendees	Location	Duration (hrs:mins)	Recorded	Transcribed/ Summarised
18	02.05.13	Info Strategy (Operations Handover)	Operations Team – Handover Planner	InfraDig, Breakout area (booth)	1:20	✔	✔
19	02.05.13	Info Strategy (CoP, RAG, next steps)	Barbara	InfraDig Breakout area (booth)	*1:00*	-	Notes
20	09.05.13	Info Strategy (Operator coordination)	Operator coordinator/ liaison manager	InfraDig, Breakout area (booth)	1:02	✔	✔
21	09.05.13	Info Strategy (catch-up)	Donald	InfraDig, Breakout area (booth)	1:22	✔	✔
22	09.05.13	Info Strategy (Design + Handover)	Asset Data Best Practices Lead (consultant)	InfraDig, Near his desk	*0:30*	-	Notes
23	06.06.13	Info Strategy (CoP)	Barbara, Mallory	InfraDig, Breakout area (booth)	1:56	✔	✔
24	03.07.13	Info Strategy (Review Workshop plan)	Donald	InfraDig, Breakout area (booth)	*1:30*	-	Notes
25	25.07.13	Info Strategy (Workshop Output)	Donald, Ralph, Mallory & Barbara	InfraDig, Meeting room	0:58	✔	✔
Subtotal					**26:28**	**11/25**	**11/11**
Other Meetings							
1	24.01.13	Introductory Meeting	Donald, Anthony, IT Architect/DBA	InfraDig, Meeting Room	0:42	✔	✔
2	16.01.13	Exploratory Call	Donald, Anthony	Conference Call	*0:30*	-	Notes
3	28.01.13	InfraDig Induction	Various new staff	Canary Wharf	3:30	-	Notes
4	29.01.13	Asset Data Academy	Software Provider - Academy Offices	London, City	4:00	-	Notes
5	28.02.13	CIO Meeting	CIO	InfraDig, CIO's office	*0:40*	-	Notes
6	04.03.13	Training links	IT Training coordinator	InfraDig, Breakout area (booth)	*0:30*	-	Notes
7	04.03.13	QlikView discussion	Anthony	InfraDig, coffee shop	*1:00*	-	Notes
8	12.03.13	CIO Meeting	CIO	InfraDig, CIO's office	*1:00*	-	Notes
9	26.03.13	Catch-up meeting (QlikView/next steps)	Anthony	InfraDig, Breakout area (booth)	0:48	✔	Notes
10	18.04.13	6 Sigma Quality Team	Head – 6 sigma team	InfraDig, Breakout area (booth)	*0:30*	-	Notes
11	19.07.13	Modelling 'traffic'	Operations - Modeller	InfraDig, At his desk	*0:48*	✔	Notes
12	05.02.13	**InfraBig**	Asset Data Programme team	BCS Offices	*0:45*	-	-
13	17.04.13	**InfraBig** (Follow-up meeting)	Asset Data Programme Business Liaison Lead	InfraBig Head Office	*1:00*	-	Notes

	Date	Meeting	Attendees	Location	Duration (hrs:mins)	Recorded	Transcribed/ Summarised
British Computer Society Data Sessions & InfraBig meetings							
14	09.05.13	**InfraBig** (Lead Doc Controller)	Document Control Manager, Barbara	InfraDig Breakout area (booth)	*0:30*	-	Notes
15	17.09.13	**Infrastructure Operator**	Asset Data Manager	BCS Offices	*0:30*	-	-
Subtotal					**16:43**	**3/15**	**1/3**
Workshop							
1	17.07.13	CoP Mapping Review & Value chain mapping	Mallory (part), Donald's alternate, James, Ralph, Barbara, Matthew (part), IT Business Analyst	InfraDig, Meeting Room	3:00	-	-
Totals					**46:11**	**14/40**	**12/14**

C.4 Artefacts Collected and (Co-)Created

	Date	Artefact	Description	(Co-)Produced	Source/ Collaborator
GoCouncil – Artefacts					
Collected					
1	22.01.13	Letter	Letter (dated 12.01.13) to Members - about New Business Model Market Insights initiative	-	Patricia
2	22.01.13	Workshop Plan (initial)	Leisure Workshop Plan Outline for discussion (Visio flow-chart structure)	-	Andrea
3	**07.02.13 + Various (5)**	**Various** MI Update Reports	Smartsheet reports tracking progress on agreed actions, circulated ahead of regular MI update meetings	-	Andrea
8	08.02.13	Workshop Plan (version 2)	Revised Leisure Workshop Plan Outline for review (Visio flow-chart structure)	-	Andrea
9	19.02.13	Leisure Services Opportunities Listing	Listing of all propositions and offers being evaluated by the Leisure team for Workshop	-	Andrea
10	02.04.13	Applications Map	1 page overview map of GoCouncil applications used (Emailed following Meeting 08.03.13)	-	Dennis (CIO)
11	26.04.13	Engine Room Details	Detailed Engine Room Proposal Pack from NBM project, fleshing out initial mock-up ideas	-	Albert
12	03.05.13	Detailed Engine Room Proposal (Revised)	As above, revised following MI discussion feedback on initial detailed pack (26.04.13)	-	Albert
13	**09.05.13**	**Various** Acorn Profiles (1 set)	Summary Profiles produced for target segments (based on Acorn Data) for workshops/wider use	-	Tanya
14	14.05.13	Community Centres Review	Business Review of Community Centres with various appendices with budget/other details	-	Alex
15	04.06.13	OtherCouncil Customer Analytics Packs	Various presentations and materials summarising OtherCouncil Customer Analytics	-	OtherCouncil (via Andrea)
16	05.06.13	GoCouncil Leisure & Culture Report	Report on the status and adequacy of Leisure facilities and capabilities (Dec 2012)	-	Alex
(Co-)Produced					
1	30.01.13	Workshop Plan Template	Shared with Andrea to aid her in workshop planning	Contributed	-
2	08.02.13	Portfolio Frame Slides	Proposed portfolio frames: Business Canvass, Question-Data (2x) & Investment-Value(2x2)	Produced	-
3	27.03.13	**Question Analysis (Leisure)**	Questions captured & summarised; 2x2 analysis performed for one proposition and next steps for follow-up discussion with Leisure team	Co-Produced	Andrea + workshop
4	27.03.13	Acorn Mapping	Excel mapping of Acorn group to Leisure propositions or offers, per workshop exercise	Co-Produced	Andrea + workshop
5	27.03.13	Revised Workshop Facilitator Pack	Revised and fleshed-out facilitator pack to run MI workshops, based on debrief of pilot.	Co-Produced	Andrea
6	12.04.13	Engine Room Ideas (Mock-up)	Straw-man ideas for a shared work space to generate momentum for wider NBM project	Co-Produced	Andrea
7	29.05.03	Enough MI checklist (Draft)	Produced to address management challenge about how much market research was adequate to assess proposals, using critical questions	Produced	-
8	29.05.13	Visual Data Mapping Ideas	Proposed preparation and presentation of MI data for Community Centre Workshop	Co-Produced	Alex
9	14.06.13	Workshop Output, & Summary (CC)	Photos capturing wall-charts and subsequent Word Summary for Community Centres (CC)	Co-Produced	Alex + workshop
10	27.09.13	**Question Analysis (Leisure + CC)**	Questions summarised with Business Canvas & 2x2 Data-Question, and analysed with summary	Produced	-
11	27.09.13	Engine Room Photos	Photos taken of the Engine Room as a record	Produced	-

	Date	Artefact	Description	(Co-)Produced	Source/ Collaborator
InfraDig – Artefacts					
Collected					
1	24.01.13	Consultant Proposal - Information Strategy	Proposed way forward on Information Strategy work for discussion with Donald (Physical copy only, with my notes)	-	Consultants (meeting with Donald)
2	28.01.13	Induction materials	InfraDig Values Handbook & Access Card	-	-
3	29.01.13	Bentley Handouts	Bentley Software induction handouts outlining key elements of the eB systems being used	-	-
4	11.02.13	Asset Information KPIs	Outline of various KPIs for discussion, presented at Project Kick-Off	-	Paul
5	14.02.13	Asset Data Dictionary Example	Powerpoint slide illustrating Data Dictionary detail for an Air Conditioner	-	Paul
6	14.02.13	Asset Reporting Requirements Workshop Slides	Pack presented at the second Asset Reporting project meeting to structure discussion. Included potential data items, discussion points and possible reporting approaches	-	Paul
7	18.02.13	InfraDig Contract Map	Visual summary of contracts against a logical map of the proposed Infrastructure	-	John
8	18.02.13	Asset Report Metrics	Spreadsheet of potential data items, indicating those available in the eB system (produced for second requirements workshop)	-	Paul
9	18.02.13	Asset Reporting Requirements Workshop 2	Slide-pack for second workshop, summarising key elements from spreadsheet analysis and flagging points for discussion	-	Paul
10	21.02.13	Photos of Work-Area	Photo of workdesk and sunset view from InfraDig offices	-	-
11	21.02.13	Asset Reporting Indicative Design Pack	Powerpoint pack outlining potential design elements for discussion at project meeting	-	Paul
12	21.02.13	Asset Report Metrics 2	Revised spreadsheet of proposed metrics, based on discussion and investigation, rated Red and Amber, for review and discussion	-	Paul
13	21.02.13	QlikView Presentation	Powerpoint slidepack setting out current and proposed use of QlikView for Reporting	-	Anthony
14	26.02.13	Asset Reporting Functional Specification	Document summarising final specification, the end deliverable for this project phase	-	Paul
15	01.03.13	Building Information Management (BIM) Working Party Strategy	Construction industry working party strategy output, including good BIM and COBIE overviews, strategies and key participants	-	CIO
16	28.03.13	Contract Design Packages (Design, Corporate &Works)	Slide detail to accompany visual Contract Overview Scope Mapping	-	Barbara
17	28.03.13	eB Knowledge Community	Draft Terms of Reference (2 pagers) listing aims, key people involved and areas of focus	-	Barbara
18	28.03.13	Quality Control Reporting **(4)**	Sample summary and detailed contract review reports focused on Document & Data Quality	-	Barbara
19	11.04.13	Document Management Action Plan	List of issues and proposed improvements for one Operator – example.	-	Barbara
20	11.04.13	Document Control Scoring	Dashboard Quality Performance score-sheet by contract (colour coded) based on two snapshots	-	Barbara
21	11.04.13	Document Control Workshop Slides	Agenda and session slides for an information-sharing session for Document Controllers	-	Barbara
22	17.04.13	BIM Virtual World	Framework linking documentation, assets, and data related elements (distinguishing Asset tags and equipment serial labels	-	Ralph
23	17.04.13	BIM overview slide	BIM elements, sequence and handover map	-	Ralph
24	17.04.13	Enterprise Change Control Process Documents **(2)**	Two documents providing an overview of the change control process and related meetings at Design, delivery and Programme levels	-	Ralph

	Date	Artefact	Description	(Co-)Produced	Source/ Collaborator
25	25.04.13	Terms of Reference (ToR) – IPAIF & AIGG **(2)**	Draft ToRs for Industry Partners Asset Information Forum (IPAIF), and Asset Information Governance Group (AIGG).	-	John
26	29.04.13	Asset Information Management Plan (V5)	**Confidential**: Overall policy document for Asset Data, outlining management structure, scope, aims, broad timeline for handover, Responsibilities and Governance (e.g. committees, reporting, etc.), supporting systems and processes	-	Mallory
27	02.05.13	Benchmarking and Performance Table	Benchmarking Data applications and contract performance summaries (RAG) across Modelling (CAD), Document control (eB), GIS & Asset Info	-	Barbara
28	24.05.13	Wall-mounted Prezi Photo	Photo of Prezi CoP-Stakeholder extracts mounted on the office wall	-	Barbara
29	06.06.13	Technical Information Principles (DRAFT)	**Confidential:** Covers ALL technical design/ engineering information, related department (including Asset Data, CAD, GIS, etc.). Includes a copy of InfraDig Information Policy (Appendix)	-	Barbara
30	06.06.13	Draft IM Roadmap with mini-PIDs	Roadmap and suggested projects slide-pack produced as output from the Information Management Strategy project (Jun-Sep 2012)	-	Mallory
31	04.07.13	InfraDig Information Management Strategy	SlideDeck, CIO signed off strategy (Sep 2012). Includes project output, maturity model, landscape, as well as vision, objectives, governance model, stakeholder map, road map and mobilisation plan	-	Mallory
32	08.07.13	QlikView Dashboard	Snapshot of Initial KPI & drill-down mock-up prepared by Consultants	-	QlikView Consultant
33	24.07.13	Visio-CRL Business Meetings Chart	One page description of various regular InfraDig meetings in a pyramid hierarchy to show their relationship	-	Ralph
(Co-)Produced					
1	04.03.13	Community of Practice Overview	Slidepack of key concepts to introduce for the Information Strategy discussions. Extended version produced on 12.03.13	Produced	-
2	02.05.13	InfraDig Heatmap Report Template	Translation of Visual InfraDig Contract map into an Excel format for flexible colour coding	Produced	-
3	02.05.13	Document Control Heatmap	Colour coded contract map to reflect quality compliance review scope and broad results	Co-Produced	Barbara
4	09.05.13	Domain Stakeholder Heatmap & Key forum inventory/Assessment	Mock-ups of two possible templates to track and assess stakeholder engagement and related forums	Produced	-
5	31.05.13	**Data CoP-Stakeholder Map**	Prezi mapping of different Data/Document Control related groups, and related meeting, ToR and codified guidance	Co-Produced	Barbara
6	08.07.13	CoP Workshop Proposal	Workshop Outline Plan to do Asset CoP mapping report back, prioritise Asset stakeholders and map out a draft value chain	Co-Produced	Barbara
7	19.07.13	Workshop Output photos (several)	Snapshots of workshop post-it mapping: stakeholder priorities and	Produced	-
8	25.07.13	Workshop Output presentation	Prezi presentation summary of workshop output and some proposed next steps	Co-Produced	Barbara

Appendix D Data Analysis

D.1 GoCouncil Workshop Question Analysis

Workshop 1 output related to exploring three Leisure service offerings in terms of potential market insights that could be sought to refine the offer and its targeting:

- Theatre in the Village (an existing service)
- School Club (new proposed service)
- Cycling Event (new proposed service)

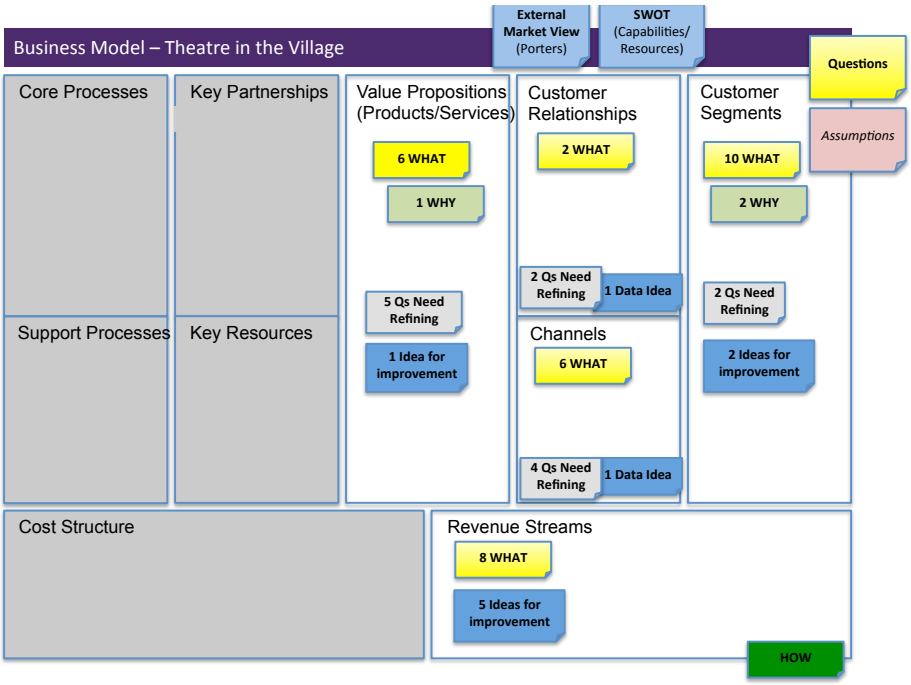
Workshop 2 output related to exploring two Community Centre options in terms of potential market insights that might be useful for deciding on which options to pursue:

- Change the Way we Operate [Community Centres]
- Transfer [Ownership or Responsibility for them]

The output and question analysis related to the first offer only, is presented below and on the following page, by way of illustration of the outputs and in-situ analysis performed.

D.1.1 Business Model Question Summary

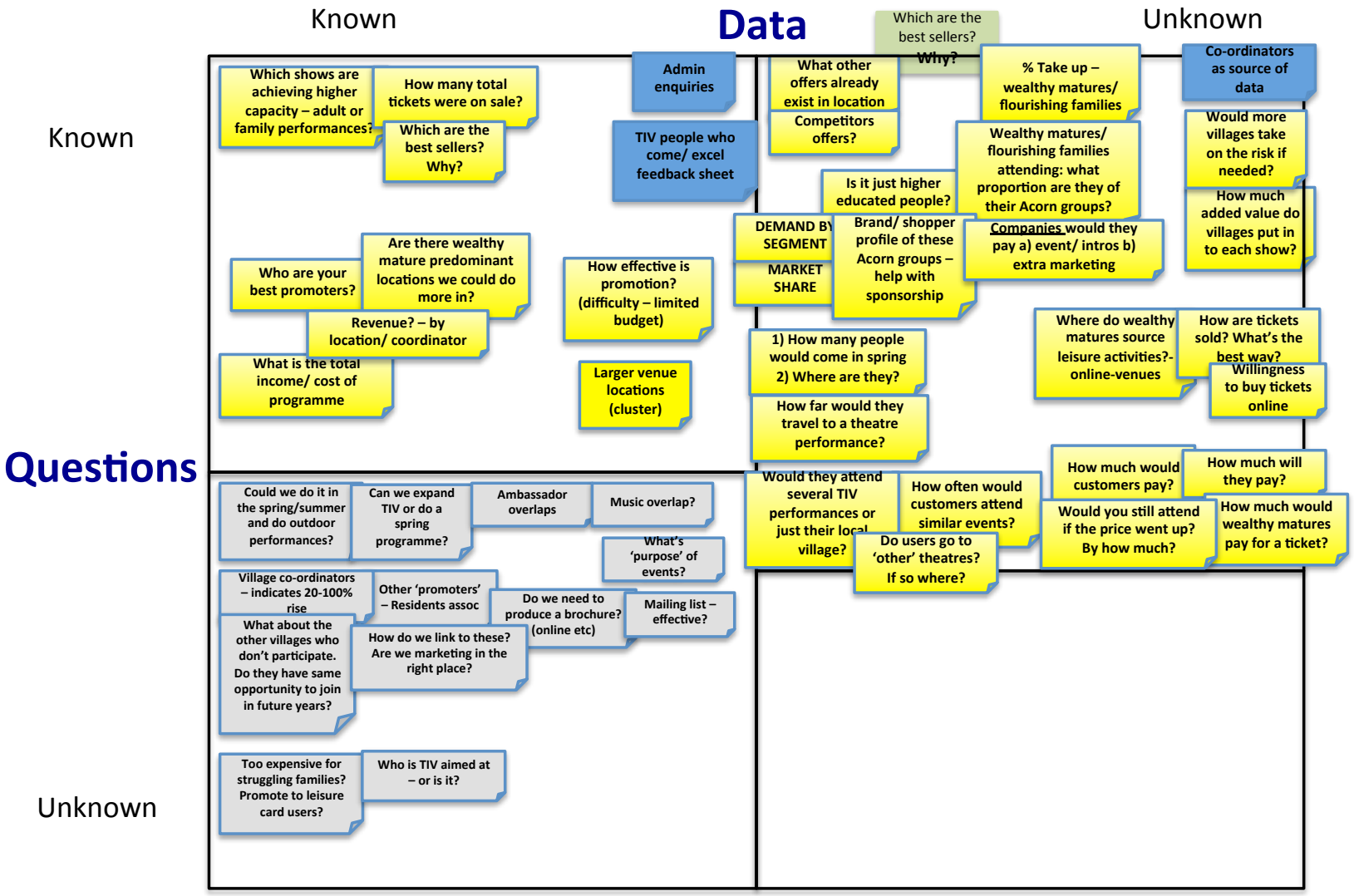
The Business Case Summary Question analysis presented first, counted the number of participant Questions captured, classifying them using Blaikie's (2007) characterisation of Research Questions as What, Why or How Questions. It also captured a number of Questions requiring further clarification, as well as ideas for improvement and potentially useful Data sources identified.



D.1.2 2x2 Portfolio Analysis (Question clarity-Data Availability)

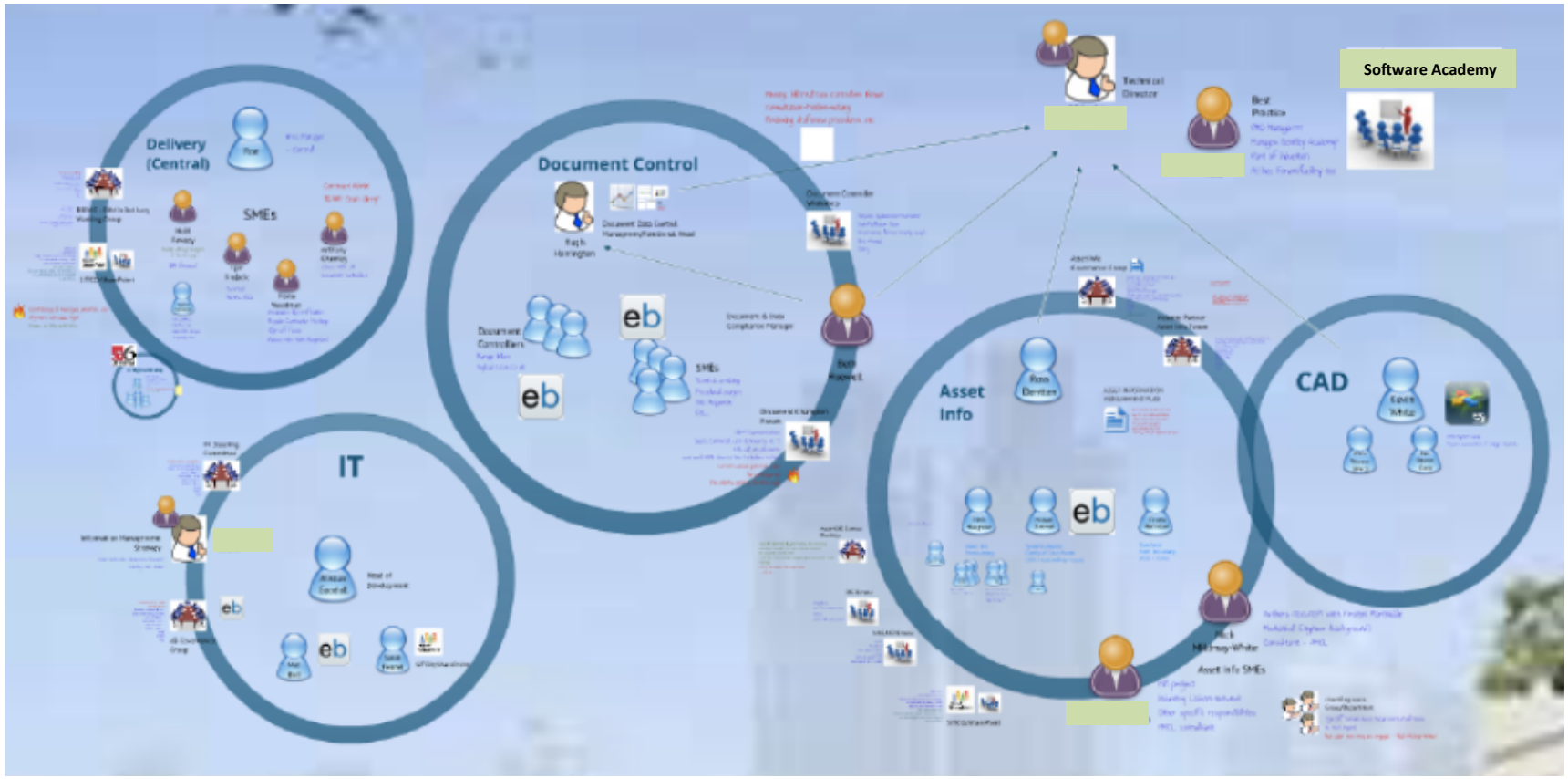
When completing the portfolio mapping of questions shown on the next page, the original post-it text and format was retained to maintain a clear visual and wording connection with what participants captured during the workshop. This was done in order to promote a sense of recognition and ownership, in order to encourage them to take forward further action, to prioritise data collection and analysis activities to address market insights thought important to improve their propositions. Some existing data sources also emerged from the workshop, which were captured in blue (for reference, as likely to be useful resources for analysis).

Portfolio Analysis – Theatre in the Village



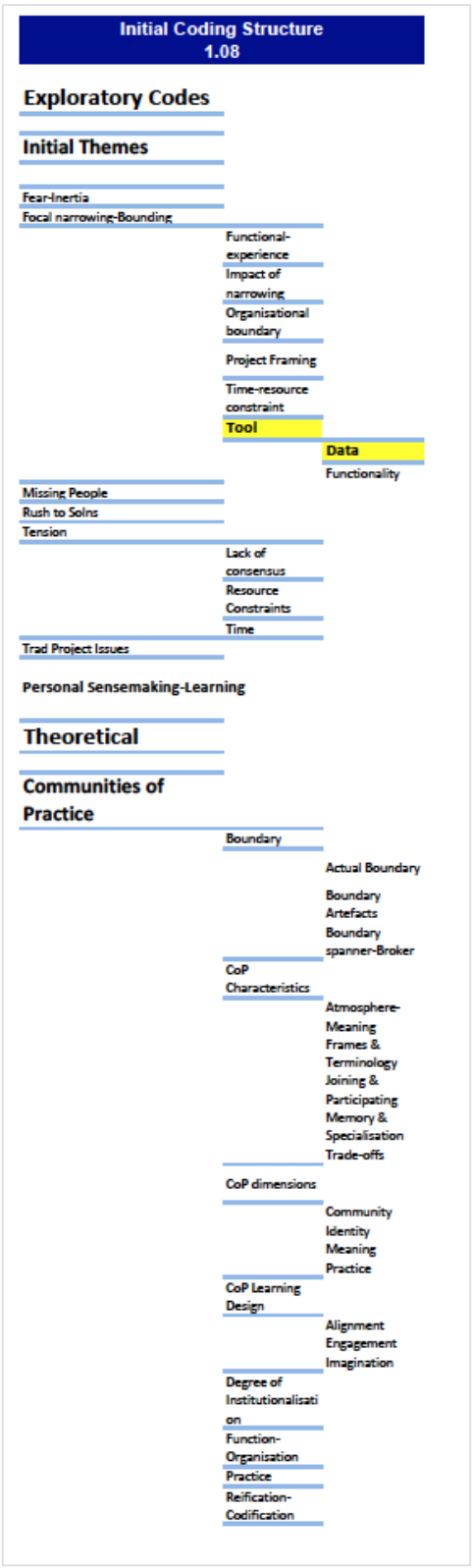
D.2 InfraDig Data related Community of Practice Mapping

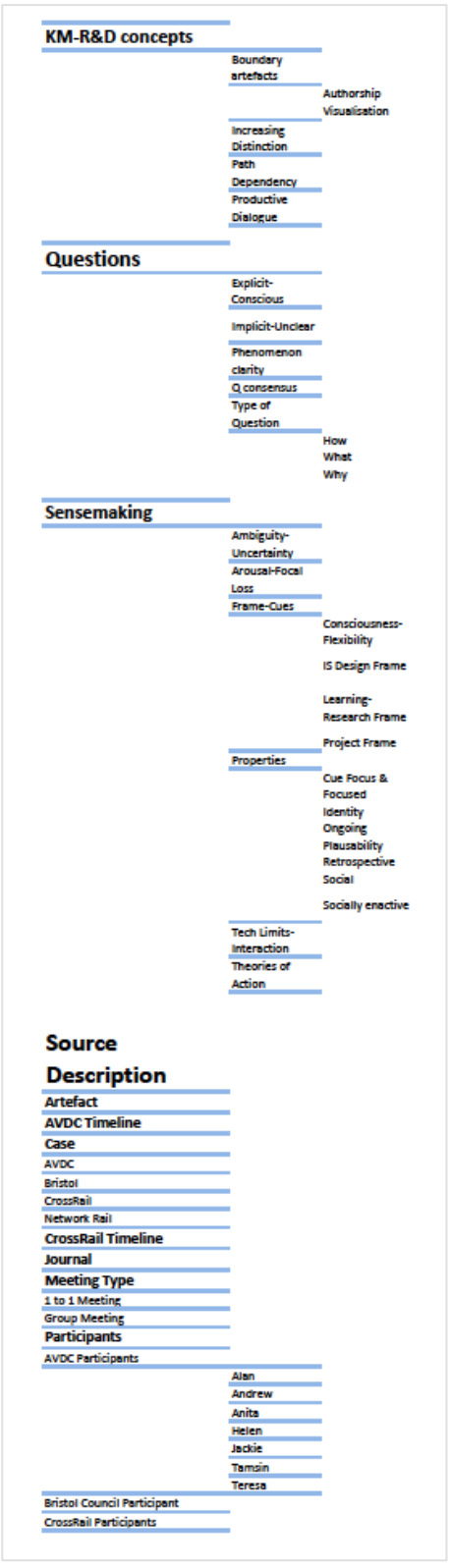
Overview of Mapping, Key and elements mapped, as well as an illustrative detail excerpt of the Head Office groups mapped


Software Academy
Delivery (Central)
SMEs
Document Control
Document Controllers
SMEs
IT
Asset Info
CAD

D.3 Coding Details

D.3.1 Initial NVivo Coding structure





D.3.2 Overview of higher density codes/elements (preliminary coding analysis)

Code Group	Higher Density Codes/Elements		
	Code	Elements	Instances
Data	■ Adequacy-limitations-missing-gaps	-	15
	■ **Issues-challenges-effort**		**28**
	■ Types (e.g. Customer)		16
	■ Orientation-Ethos		16
	■ Usefulness-Purpose-Significance		15
Initial Themes	■ Fear-Inertia	■ Effort to engage	14
		■ Impact concern	13 +10
	■ **Focal narrowing-Bounding**	■ **Functional-Experience**	**20**
		■ Prioritisation	13
		■ Time-Constraint	15
	■ Missing People	(especially IT representation)	14
	■ Tension	■ Conflicting personal objectives-factors	11
		■ Lack of consensus	12
		■ Resource Constraints	13
	■ Urgency-Time pressure	-	9
Communities of Practice	■ Boundary	■ Boundary Artefacts	10
		■ Boundary Spanner-Broker	14
	■ CoP characteristics	■ Frames & Terminology	11
	■ **CoP Learning Design**	■ **Engagement**	**23**
	■ **Function-Organisation**	■ **Silo Issue**	**21**
KM-R&D concepts	■ Boundary-Sensemaking artefacts-aids	■ 2x2 Questions-Data	11
		■ Business Canvass Framework	11
		■ Visualisation	13
Questions	■ **Explicit-Conscious**	-	**21**
	■ Implicit-Unclear	-	11
	■ **Phenomenon Clarity**	■ **Customers**	**23**
		■ **Services**	**17**
	■ Type of Question	■ What	16
Sensemaking	■ **Ambiguity-Uncertainty**	■ Clarity of Purpose	11
		■ **How to proceed**	**17**
	■ **Assumptions**	-	**19**
	■ **Frames-Cues**	■ **Customer Orientation**	**20**
		■ **Financial-Commercial Frame**	**28**
		■ Political-Power-Buy-in Frame	14
	■ Personal Sensemaking Learning	(me)	16
	■ Theories of Action-Constraints	-	14
Tool	■ Integration challenges	-	7
	■ Tool Type	■ CRM	9

D.4 Participant Interview Summaries (extract)

Participant Interviews

		GoCouncil			
Interview Question	**Phenomenon Aspect Covered**	Tanya	Albert	Andrea	Helga
What is your understanding of the Initiative underway, in terms of its scope and overall objective(s)	**Framing of Project Purpose**	'We want to sell stuff' 'We need to be more like Amazon' (suggestions/cross-selling)	'Getting to a balanced budget' 'Cutting our cloth', 'Burning platform' Cutting costs & Growing Income (dual objectives) Related Culture Change - Need driven/ more commercial approach, 'turn a tank' Have a conversation with residents about what council tax covers (a third)	'Re-orienting the council' 'meeting the [Funding] shortfall' 'What kind of council do they want to be? & How [to get there]?' New Business Model capability (trained, engaged and delivering) Clear Vision and enabling Management Culture	Become self-supporting by 2020 Way of redefining what we do Stopping stuff customers don't want Start doing what they want & will pay for (profitably)
What phenomenon are you trying to find out more about, which is at the heart of the initiative?	**Framing of Phenomenon**	Need to know more of customers, residents, future customers	Customer insight - become clearer - we don't understand what customers want or know our customers/residents	Don't understand/know their customer (currently what they think they need)	Market Insights to inform the New Business Model process and test assumptions, related to Markets, Customer Segments & Access routes to them Different perspectives (esp. Customer's)
What are the dimensions or aspects about this you are most interested in understanding better? OR what ideas about it do you want to test or confirm?		Not addressed	Not addressed	Not addressed	Not addressed
How would you frame these as questions you are seeking answers to?	**Questions**	What to sell?	Data will prove for Residents... 1. [Are there] some [services] they will pay for or want more of? 2. others they don't value? 3. what are key social core free services they are happy to cross-subsidise?	What kind of Council do they want to be? & How [to get there]?'	New Business Model: Where are the best Markets? Which customer segments/who is interested? Best ways to target them? Building Control: (specifically) Who is doing building work? What would they be prepared to pay? [for a premium service]
	Project Concerns/Challenges	Silo behaviour and Data (eg lack of alignment) Lack of Data Visibility, Ethos & Engagement Assumption-driven behaviour Acorn Data Quality and limits - eg Average data, postcodes Data-a part-time job, no home, ownership	Data will only get you so far' (given the socio-political overlay) Public-service rather than commercial Ethos 'Party Political overlay' & election cycles 'Is the Burning Platform hot enough' (given history of finding fairly painless cuts) Is the New Business Model the right way to go? Elephants in the room - not for discussion (High profile-political projects, Retired Age profile) Willingness of managers to take tough decisions Data - friend or enemy? (at it's mercy and in light of emotional rather than rational responses)	Many different views of WHAT they want? (CEO clear but not a clearly articulated Vision, & consensus outstanding) HOW is not clear (& greater challenge) Silo, functional working (not leveraging limited business skills & experience, Data, synergies) Cultural Shift (from Rich DC to constrained, need to demo value) Propositions based on assumed need currently (Assumption based - don't unpack them) Don't measure anything (ie not data-driven, data supports/reinforces decisions/outcomes) Access to data is difficult CEO lack of understanding & frustration - Just Do It approach Internal Communication & Coaching/support (new skills/frames required, us and them) Managers a block, not on board & enabling	Silo Working (web example) Lack of communication - managers filtering/blocking Dysfunctional team-politics-existing problems Historical service focus rather than customer/value focus and mindset 'chasing a ball' - leader attention led Lack of 'speed' or urgency/delivery

Participant Interviews

		InfraDig				
Interview Question	**Phenomenon Aspect Covered**	Project Engineer Contractor	IT Business Analyst	**Mallory** Chief Engineer	**Ralph** Asset Data Team Lead	**Anthony** IT Systems Development Manager
What is your understanding of the Initiative underway, in terms of its scope and overall objective(s)	**Framing of Project Purpose**	Formal (periodic) reporting **mainly** and ad hoc (day to day) reporting Little apparent knowledge/appreciation of Asset Info Initiative (mentions EB/Asset Info only as an afterthought when describing key systems)	**Objectives:** Broadly - Consistency of Master Data) - data quality (discrepencies), to enable 'slicing and dicing, Reporting focus (Performance) KPI project objective seen as making sure contractors are capturing/updating virtual asset data to avoid completeness issues in due course at hondover (5 yrs time) **Scope:** Focused on Core systems (Planning, Financial/Cost, Risk, H&S) Not really Asset/Construction side/systems - currently seen as out of scope although recognises likely to be overlaps/linking with Finance (assets).	Keen to transform maintenance & operations Standardisation to enable collaboration Wants owners (in particular) to see virtual and physical asset as equally important	***KPI project*** *Didn't really address this* **Wider Virtual Asset** Framed as Master Data Management to: Provide Single source of data type and Remove duplication of data sets *Does recognise wider establishment, maintenance, optimising and handover per role description above too*	**KPI project** - to 'police' the data within EB by providing a view of what's there, i.e. Reporting **Wider Virtual Asset** - What is in EB, will be handed over (seen as simple). Wider piece really for [chief engineer]
What phenomenon are you trying to find out more about, which is at the heart of the initiative?	**Framing of Phenomenon**	Effective delivery/program control (main focus) Not really engaged on Asset Info side - seen as technical (Malcolm's area) - see below though	*Performance?* *Not clearly articulated*	*Not clear he has a learning objective* *Rather has a change objective and believes better/integrated data is the enabler*	*Purpose provided rather than Phenomenon but implied as:* *Inefficiencies in the way vital items are constructed (1) and maintained (2)*	*Implicit:* *EB Asset Data compliance/DQ 'police'* *Personally - interested in resolving IT issues*
What are the dimensions or aspects about this you are most interested in understanding better? OR what ideas about it do you want to test or confirm?		Workbreakdown structure seen as key: Contract, Project, Location, Activity/Activity Group Beyond that various dimensions, depending on focus. Criticality comes up (not explained)	Not addressed	Not addressed	Not addressed	*Performance & Discovery aspects* *Not really clear on the detail behind these though (sees these as Asset team's area of expertise)* *Mentions contracts* *IT issues data: people, location/site issues (much clearer - closer to these)*
How would you frame these as questions you are seeking answers to?	**Questions**	Sees many as emergent, especially for Programme Control. Inferred things like... What project issues are emerging requiring intervention? Contracts that are late? Why/detailed progress to resolve, etc? For Asset Data, sees key Question as: How am I going to Run & Maintain this asset?	Not addressed	*Not specifically addressed but implied:* *What inefficiencies can we identify in Asset maintenance?*	*Not specifically addressed but implied:* *What inefficiencies can we identify in Asset construction (1) and maintenance (2)?*	*Recognises Question-led - Illustrative Questions provided:* *Well, why are they doing so much better than them? [contractors]* *Where should contracts be versus where are they?*
	Project Concerns/Challenges	Siloed working & lack of data integration Lack of Management commitment to integrate IT delays - 18 weeks to provide additional data Business ownership of systems (even infrastructure). Surfaces/called 'politics' Asset Info - doesn't believe they've thought about likey use 'what are you going to do with all this data' (Clear Purpose) - more collection focused Asset Info - different systems at [Operators] -seen as a critical area to align to get to use A lot of New Staff coming on board that need to get up to speed, aligned, etc. Visibility and Communication seen as key, especially as rapid learning occurs	Information/Data handover seen as a key current difficulty (that Asset Info will address) Contractor buy-in to providing asset data Completeness of Asset Data Coordination across teams, especially if discontinuous timing (of involvement, eg. Across phases/contracts, etc). Getting alignment & agreement Inconsistent Master data - e.g. Location codes example (also seen as a Key dbase link) Lack of visibility/understanding of other siloes (IT seen as having more of an overview but lacking detailed knowledge of each area- complex)	Contractors - getting them to change/do data Lots of things to integrate - time required Tool-data immaturity & incompatabilities (DGN files vs CAD DWG formats) (Frustration & impatience/expectation) Focus on People as particularly important and no-one is focusing on them - Biggest challenge (People Process Technology cited -sees People as simple, Technology as straightforward) Getting engagement seen as key-fun/skunkworks Frustration with IT (delay) Client focus on Total Cost of Ownership and prescriptive requirements (alignment & detailed commitment-'enshrined...', lip service & lack of understanding, complex) Lack/inadequate training & bandwidth (day jobs)	Establishing and Maintaining the integrity of the data The cultural aspects of implementing good data management & control practices Contract levels identified (3) with Design and Build ideally together but sometimes split. Design lags noted when construction pushes back on/changes design - this is where he sees change control as key to ensure appropriate impact assessment, and subsequent audit trail. Saw change control as new to rail construction Seemed to have a clear distinction between Project Information Management (PIM) and Asset Information Management (AIM), with the former overlapping the latter but stopping at handover, while the latter continued with the Operator/Maintainer (seems to have had experience of this aspect)	Project and Requirement specification seen as problematic, causing delays - 'guerilla/BAU approach prefered' Difficulty/challenge defining data requirements (especially for 2nd order/follow-up Qs) Consistent definitions seen as NB ('taxonomy') - endless meeting/discussion frustration (then doesn't work technically anyway) 'Too many fingers in the pie' - all with a budget to spend on 'toys', resulting in loss of focus/results IT brings too much overhead/process/cost 4 skunkworks (believes should be business led) Not much confidence in KPI report project outcomes IT function/accountability overview Complex (partially devolved) IS approach confuses but ok given short term nature [of InfraDig] ClickView sensor project overview for building movement a data volume challenge

Participant Interviews

		InfraDig (continued)				
Interview Question	**Phenomenon Aspect Covered**	**John** Asset Data Team Member	**Paul** KPI Project Consultant	IT Architect/DBA	**James** Asset Data Consultant	**Matthew** Asset Data Super User
What is your understanding of the Initiative underway, in terms of its scope and overall objective(s)	**Framing of Project Purpose**	**KPI project** - to provide a tool to understand completness, consistency/quality and timely capture of asset data. To be able to identify holes **Wider project**: Provide operator with 'fit' information to operate and maintain the asset	**KPI Project** - seen as similar to CAD and Document Control Projects. Provide Indicators of progress by contract, project, area bases, to highlight issues and allow intervention early Trying to automate reporting as much as possible to avoid errors & chasing contractors (sees granular Asset reporting as out of scope?) **Wider Project**: *Didn't really address this (seemed reticient to volunteer views here)*	Not covered directly *Doesn't seem close/engaged with these projects* *More reactive-oriented to business requests/ tasks/projects as 'cascaded' by [IT systems development manager/ IM strategy manager]*	**Wider Project:** Framed as Data maturity 'Journey' Already progress - initially no clarity, now Governance, DD, AIMS in place Early data collection as key Realisation that virtual asset as important as physical Enabling continuous improvement Handing over Asset Data to Information Managers [Operators] **KPI Project:** Keeping on 'haystack' of Asset data - won't fall over & no 'needles', just 'all good hay' + Provide tools to monitor progress, process and fitness/quality of data gathering + Transparency (status) - contractor peer pressure Iterative/refined requirements will emerge	**KPI Project:** A tool to provide visibility, mainly to senior Managers (Big Picture) & Middle Managers (Project Managers) Believes instead in building incremental use and buy-in through engagement/demos/problem solving using asset data (mockups/visualisations) **Wider Project:** Not very clearly articulated... 'Virtual Railway' but not all the stuff they want' Everything linked/related but not fluidity/ease of use/in one step Sees progress over previous project, virtual emphasis over paper, earlier data clarity/ collection Requires more iterations 'to get it right' ***(use** of IT/Data in construction? not clarified)* Learning curve as to what Asset Data can do, anticipates 'getting some way there' Not just 'ticking a box for BIM' (as on prior project) Earlier Data than previously so more learning likely
What phenomenon are you trying to find out more about, which is at the heart of the initiative?	**Framing of Phenomenon**	Contractor performance in terms of data provision	**U**nderstand progess on populating asset data and validation-signoff (different levels for different users/audiences - board, PMs, Asset)	Not covered *Not clear he is close to this*	*More interested in how to effect change/mindsets than in underlying Asset Data & Maintenance phenomenon (perhaps given experience feels he knows the answer?)*	*Not very clear on this beyond Asset data problem solving, how to encourage Asset data use [likely to be mainly project oriented rather than downstream maintenance]*
What are the dimensions or aspects about this you are most interested in understanding better? OR what ideas about it do you want to test or confirm?		Completeness, consistency, quality, timeliness Comparison of contractors at different stages (rather than just at a given date) Top Down and Bottom up views	Progress by contract, project area, and overall (depending on audience) Validation mentioned once *(possibly reflects less emphasis on Quality dimension at this stage, eg. no dimension around owner/signoff accountability - [Operators] or data type/class mentioned. Possibly seen as detailed)*	*Only engaged to the extent this is reflected in current cubes, database feeds, etc.* *Particularly interested in potential Links/Keys.* *Sees requirements as emerging over time, prefererence for providing everything initially*	*Data Quality, Progress on Asset Data capture, consistency*	Don't know what detail will be required at this stage - 'guessing'- likely to emerge (KPI project)
How would you frame these as questions you are seeking answers to?	**Questions**	Not really clear/articulated at this stage beyond contractor comparative performance, identtifying holes	*Do I need to intervene to address asset data?* *Which contracts?, which data?* *Why is this area/contract behind?*	Not covered *Not clear he is close to this*	Are there any 'needles' in my haystack or all good hay? How well am I doing compared to other contractors? [transparency of data capture, peer pressure]	Not covered
	Project Concerns/Challenges	Sign off Data requirements by [Operators] (attributes to collect)- 'raising the bar' for them Different Coding taxonomies (esp. Civils) - getting agreement/standardisation or mapping Getting people to change (esp. [Operator] entrenched ways of working) Recruiting enough appropriate resources to onboard before ramp up in data collection key KPIs - Getting Data at the right level of detail for different audiences (board versus day to day) How to use KPI data effectively to highlight the right messages/focus (key issues, etc.) Volume of data collection in due course and getting stakeholder signoffs (esp. external [Operators] Contractors - familiarity with Data requirements, processes, hidden tiers (i.e. subcontractors), consistency/complexity Steep learning curve (himself) and new guys too	Key dependency for success is use of KPI data to identify issues and intervene early Based on CAD experience, getting data from Contractors seen as key challenge Level of contractor IT kowledge/savvy seen as lower than usual (construction sector) (will need more handholding) *Mentions 'crash course' in Asset Data, although not as a challenge specifically*	No obvious architecture issues impacting on KPI/wider project identified/seen, **unless** a move to real-time or daily 'snapshot' requirement. Currently monthly period load (entire dbase) Main impact seen as likely to be on ETL (Extract Transform Load) side i.t.o. data volume growth More generally: - No real documentation of current architecture, hosting arrangements fragmented, accountability, configuration decisions, etc. *(seems to have evolved...)* - Some single dbase sources - EB currently (setting up a replicated version) - EB data & fields, but no structure documentn. (access to developers though) - No real optimisation so far - starting to target performance issues tho (example)	PMgr/Contractor mindset change about importance of virtual data asset (contractual, lack of goal alignment/ incentives, physical bias)-not just snag-list item Culture/behaviour change challenge - seen as manufacturing in the 50s - Quality control versus anyone can stop production line (pervasive value/action) - Downstream impact rather than immediate feedback (typically on someone else/goal alignment issues too) (Underspend/ineffective maintenance only emerges over l/term) Misunderstandings/lack of clarity/consistency (e.g. requirements) [Operator] legacy system/data - Elipse (versus [Another Operator]), which has a relatively clean slate). Data migration/ use-integration challenge =>provide in a usable format [Operator] haven't implemented meta-data for document management (mapping filenames required) Different classification approaches (historic)/ policy-system inconsistencies. Need for standardisation/convergence seen as key (pragmatic problem-solving versus theoretical) *Some political 'people' issues **(sensitivity)*** Data seen as important but not urgent, which it *is* for project to avoid rush/poor Data Quality IT beaurocracy-budget holder rather than business ownership/driven for new initiatives/projects	KPI - Balancing differtent visions between asset data and engineering teams (practice leadership vs global/broader) Wider projects -time constraint to do everything envisaged (lag between thinking & design, engagement/iterations) Complexity/detail needed [for Asset team] may not be possible in KPI project. Uncertainty as to requirement at this stage ('guessing', likely to emerge, ad hoc) Wider: Steep learning curve (what is involved and how data can be used) Communication and visibility - currently very Corporate 'press release' focused, vs middle/lower staff level or influencing - e.g. forums/channels (across silos/teams) Need for multi-level engagement - [Operators] to get practice led engagement/use - internally across teams to ensure common vision and understanding, building out straw-man solution from the bottom up (problem solving at the coal-face, sharing too) - across silos (Business/IT, other units too) -Ownership and engagement link Tool inflexibility - design/vision/changes Lack of data knowledge (generally), hence engagement and visulation to encourage this. Also skunkworks to build a 'fan-base' based on solving practical (emerging) problems & use *(quite project/construction focused?)* Hearts & Minds/bridge generational IT issue Fear/reluctance about floating new ideas (not fully

D.5 Initial comparative thematic analysis

Strand of Analysis	Question Analysis	Themes per Diary Reflection		Coding	Creative Narrative Analysis
Case Coverage	GoCouncil only	GoCouncil and InfraDig		GoCouncil only (**Incomplete**)	
Iteration-timing	February-March and October 2013 (1st and 2nd workshop outputs)	Mid-April 2013 (Supervisor – 'Very messy')	July-August 2013 (Colloquium/Supervisor discussion/Coding themes)	September-October 2013	October 2013
Questions	■ Existing services generate more questions than new offers, so prior knowledge may be important ■ Almost all questions represented WHAT questions, and many questions require further refinement/specificity, both pointing to unfamiliarity ■ Many questions could be clustered and categorised or themed: service needs, pricing, channel, which could facilitate broader, coordinated market insights research across propositions and functions ■ Volume of questions requires prioritisation	■ *Prioritisation challenge in terms of volume of questions/potential lines of inquiry*	■ Clarity about the phenomenon and related questions is often implicit rather than explicit, with an overriding focus instead on action and outcomes rather grounding them in supporting evidence, or testing underlying theories of action ■ The volume of questions a new phenomenon generates can be a significant challenge, especially to achieve consensus on which lines of inquiry to prioritise and pursue	■ A mix of many explicit and implicit (mostly WHAT) Questions were noted, especially early in the project and when prompted in workshops ■ Phenomena, underlying assumptions and related theory were much less explicit and lacked clarity, e.g. AVDC were really interested in the intersection of Services and Customer insight rather than just Customer insight ■ Questions seem to be used as a device to engage with the phenomena, explore and draw attention to aspects of it	
Data	■ A surprising lack of internal service data emerged (and lack of clarity of what data exists) ■ Acorn considered to have limited value, generally as it was average data for post-code areas, and where propositions involved channel intermediaries, rather than a direct service to households	■ *Limits of Acorn granularity (household level, averages)* ■ *Internal, (existing) tool-centered focus evident in KPI and Clickview projects* ■ *Tool frames data available/ considered (eB-KPI project) (incidentally (pre-)defined)* ■ *Idea of data as an enabler of exploration rather than just answer to a specific question.* ■ *Also notion it needs to evolve*	■ Exploration (including R&D and experimentation) for new data and phenomena, versus Exploitation approaches where data is available, known, structured, and underlying relationships-theory understood ■ Phenomenon coming into focus as important – social versus physical (e.g. infrastructure, location) is likely to require different data (easier-more difficult to define, measure and different analytical techniques likely) ■ See also below – under Focal narrowing-Bounding (Frames)	■ While data gaps and adequacy concerns were raised, the main challenge noted regarded the level of effort required to engage with data properly, both in terms of practical logistics preparing and organising it, as well as trying to make sense of it ■ A supportive ethos that is data centric (or not in AVDC's case) and a recognition of its purpose and usefulness, were seen as key to provide support for the significant data engagement effort required	■ Draw attention to Data and Tools and the role they play, especially constraints (e.g. eB data model, incompatability across user organisations, consensus on data required) ■ Tacit knowledge presupposed/required and experience/practice-testing to use data effectively

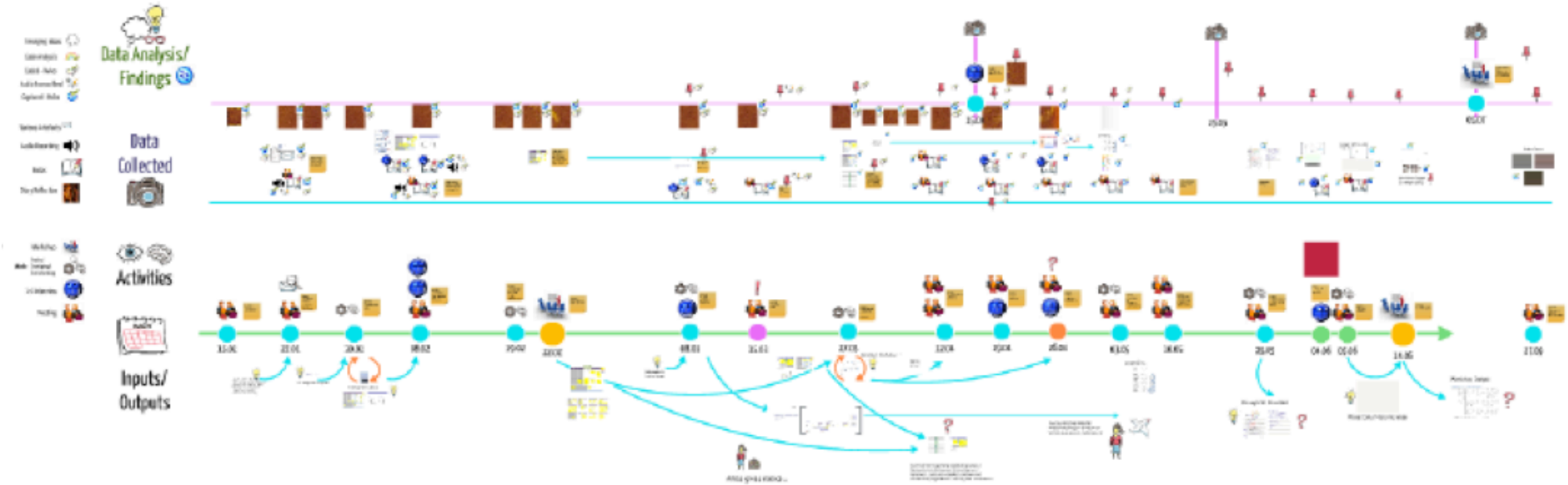
Strand of Analysis	Question Analysis	CoP-Function and Dependency 'Mapping'	Themes per Diary Reflection		Coding	Creative Narrative Analysis
Tools Rush to (optimized) solutions			■ *Rush to (optimized) solutions* ■ *Bias to existing tools and solutions (e.g. eB), which in turn frames data and approaches* ■ *E.g. Clickview and pattern focus when requirement is simply variance analysis/trends and volume/timeliness is the real challenge. Simple GIS visualization would work (see above too)* ■ *eB Tool/Data and performance constraints*	■ A rush to solutions is noted, particularly by IT function, typically seeking to optimize too early and not allowing users to experiment, clarify and evolve requirements through prototypes and engagement ■ See also below – Focal narrowing-Bounding (Frames)	■ Data and tool integration challenges were raised but the initiative at AVDC was not predicated on the use of tools ■ The absence of (and potential need for) an existing CRM system was raised ■ IT were not really involved (AVDC)	■ Draw attention to Data and Tools and the role they play, e.g. CRM fixation
Frames/ Purpose Focal narrowing/ Bounding	■ The Business Model Canvas proved an easy to understand framework for generating and categorizing market insight questions, and also to facilitate discussion and gain consensus on issues identified, although it didn't highlight competitors ■ In fact, participants struggled with the concept of competitors and to generate related questions, pointing to a possible lack of experience of competitors		■ *Purpose, Vision and scope not clear for IM strategy* ■ *People challenge recognised but not 'unpacked' (IM strategy)* ■ *Silo functional perspectives and experience bound framing about purpose, relevant data and tools, drives behavior (e.g. priorities, data sharing, problem solving)* ■ *Sustainability objective per BIM working party document not evident or talked about (KPI/IM projects). Lack of a champion?* ■ *Lack of a data-centric frame or person* ■ *Idea of project frame as problematic for initiative (rather than human-centric learning frame)*	**Focal narrowing-Bounding** ■ **Functional-experience** seems to frame phenomena and questions, limiting what participants see as relevant, especially if there is a settled, dominant or codified view, often grounded in key implicit assumptions ■ **Impact of narrowing** is anticipated in the form of missing questions and data, undermining understanding ■ **Organisational boundaries** act as barriers to communication and joint learning and action (e.g. silo thinking and behavior), especially at mid and lower levels of the organization, where fewer forums exist, but also top-down (e.g. vision artefacts) ■ **Project Framing** focuses attention on iron triangle dimensions and People-Data use aspects are missed - i.e. not managed or monitored ■ **Time-resource constraints** seem to constrain ambition and the level of data engagement/learning possible, especially when (senior) urgency or pressure is added (as this takes longer than expected) ■ **Tool/Data experience** seems to impact framing, through datafields available within existing tools	■ Attention and prioritization most often seemed to be grounded in functional perspectives and experience, with some evidence for potential narrowing or bounding (especially under time or resource constraints) ■ Discussion most often reflected a focus on Customer orientation and a financial commercial framing of issues, with sensitivity to questions of senior management and political buy-in and support ■ The commercial framing was also evidenced in the use of visual artefacts, especially the Business Canvas Framework, to facilitate engagement and sensemaking ■ Several underlying assumptions and theories as to action and constraints were noted, which were typically left tacit ■ Senior management attention seemed to focus on physical artefacts or outputs rather than learning outcomes	■ Lack of clarity of purpose, roles, How and consensus (AVDC and InfraDig IM strategy) ■ Resulting inertia ■ IT project framing ■ Organisation boundary constraints and silo behavior/thinking ■ Draw attention to artefacts ■ Draw attention to Frames and connect them to the background/ experience they arise from ■ Highlight constraints (especially time and resrources) ■ Impatience and urgency ■ Draw attention to people dimension/forums ■ Theory and assumptions underpinning value creation ■ Bias for hard artifact outputs over softer outcomes, missing in handover plan

Strand of Analysis	Question Analysis	CoP-Function and Dependency 'Mapping'	Themes per Diary Reflection		Coding	Creative Narrative Analysis
Practice community/ Situations/ Codified Process/ Sensemaking Artefacts			■ *Silo functional perspectives and experience bound framing about purpose, relevant data and tools, drives behavior (e.g. priorities, data sharing, problem solving)* ■ *Tension with IT involvement - role as driver of project rather than facilitator, solution focused/bias* ■ *KPI project very traditional waterfall project process and artifact driven (e.g. focus on requirements definition project) rather than interim progress (e.g. 'skunkwork' prototypes as interim, evolving solutions). Trade-off recognised as to when to optimize (not too early)* ■ *Wider, traditional project issues also noted, impacting on progress*	■ **Traditional project issues** noted that remain important on these projects (although not the main focus for this study) ■ **Tension** both at a personal and team or cross team level. Typically arising from a lack of consensus, especially on how to proceed, and what to prioritise in the face of resources and time constraints ■ See Boundary framing and silo issue already highlighted above ■ Sensemaking artefacts, particularly those adopting visualization approaches seem to enhance engagement with data, especially for inexperienced or cross-functional personnel (e.g. CoP mapping, value dependency)	■ Functional boundaries emerged as important, driving in silo behaviour, hampering communication, coordination, data and resource sharing, and necessary learning ■ Practioner functional experience seems very important in framing the initiative, what data/tools are relevant and how it should proceed ■ Considerable use of marketing terminology was evident (as well as some public sector specific terminology) ■ The marketing team seemed to be seeking and playing a boundary spanning, facilitation role	■ Power and Social Drivers ■ Hiccup – how much is enough/urgency ■ Silos and bounded thinking (e.g. missing IT/Data understanding ■ Include and draw attention to artefacts and their role in achieving engagement ■ Project meetings ■ Workshops ■ Data immaturity?

D.6 Case Timeline Analysis

D.6.1 GoCouncil Timeline Overview

Such a mapping and analysis was produced for each case as an aid to synthesis and precursor to producing thick case descriptions and monograms.

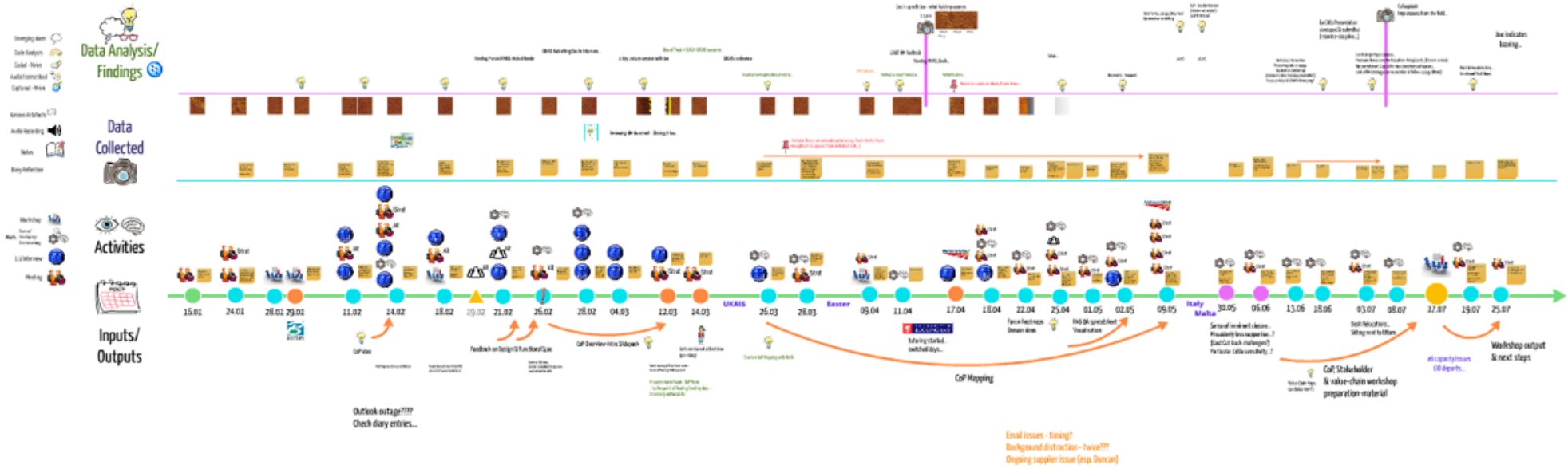


Notes

1. The timeline is shown as the Green line, with activities outlined just above, and inputs and outputs shown below the timeline.
   Arrows indicate links and differentiated colours and dot sizes highlight pivotal meetings identified
2. Above this are two further sections:
   a. the middle section summarises the data collected, its capture into NVivo, as well as transcription and analysis performed to date (i.e. the status of detailed analysis);
   b. the top line summarises diary reflections as well as highlighting particular points in time when an informal review was undertaken and key themes were informally identified (as a record of the sense I was making at the time).

D.6.2 InfraDig Timeline Overview

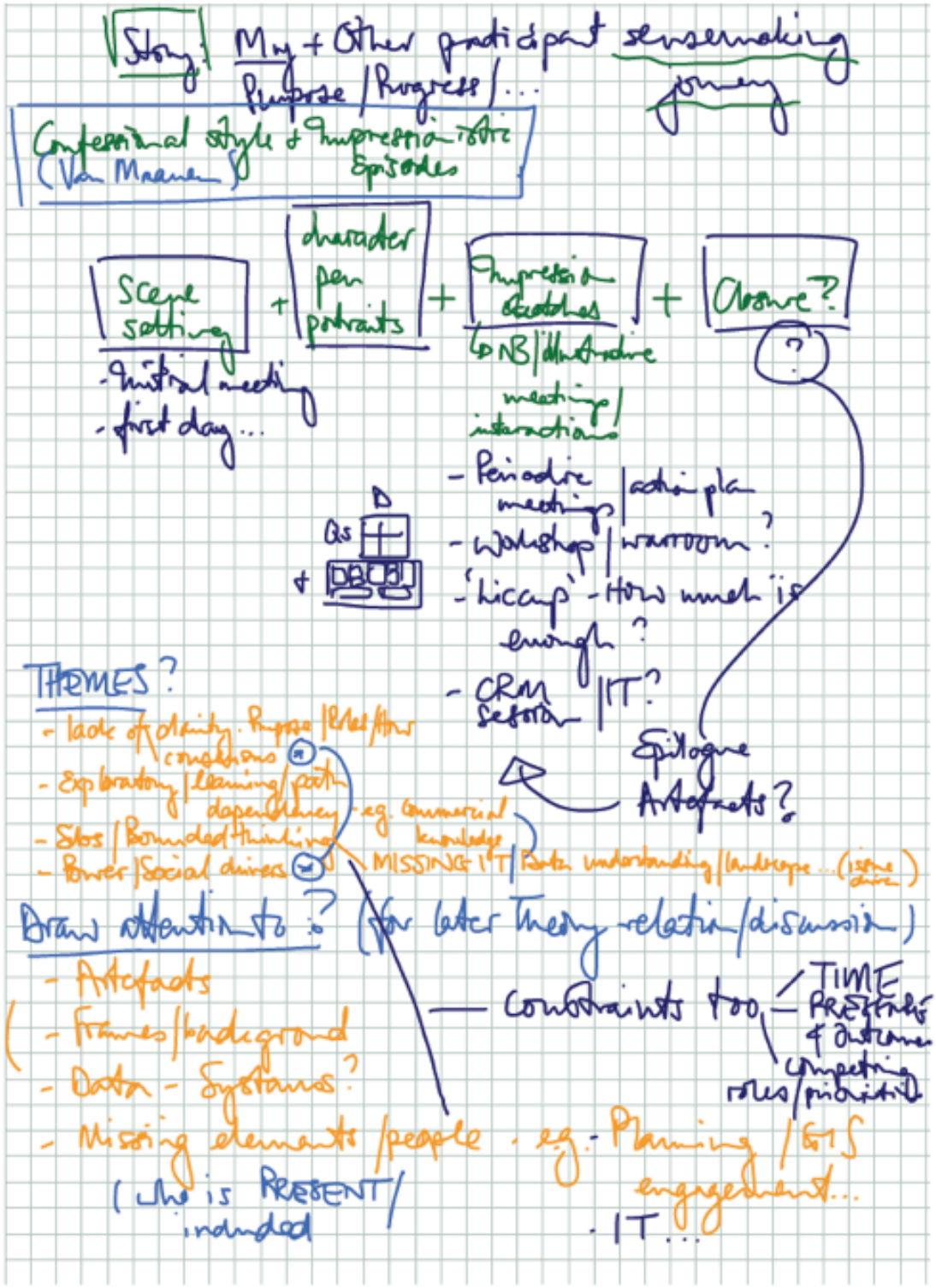
Such a mapping and analysis was produced for each case as an aid to synthesis and precursor to producing thick case descriptions and monograms.



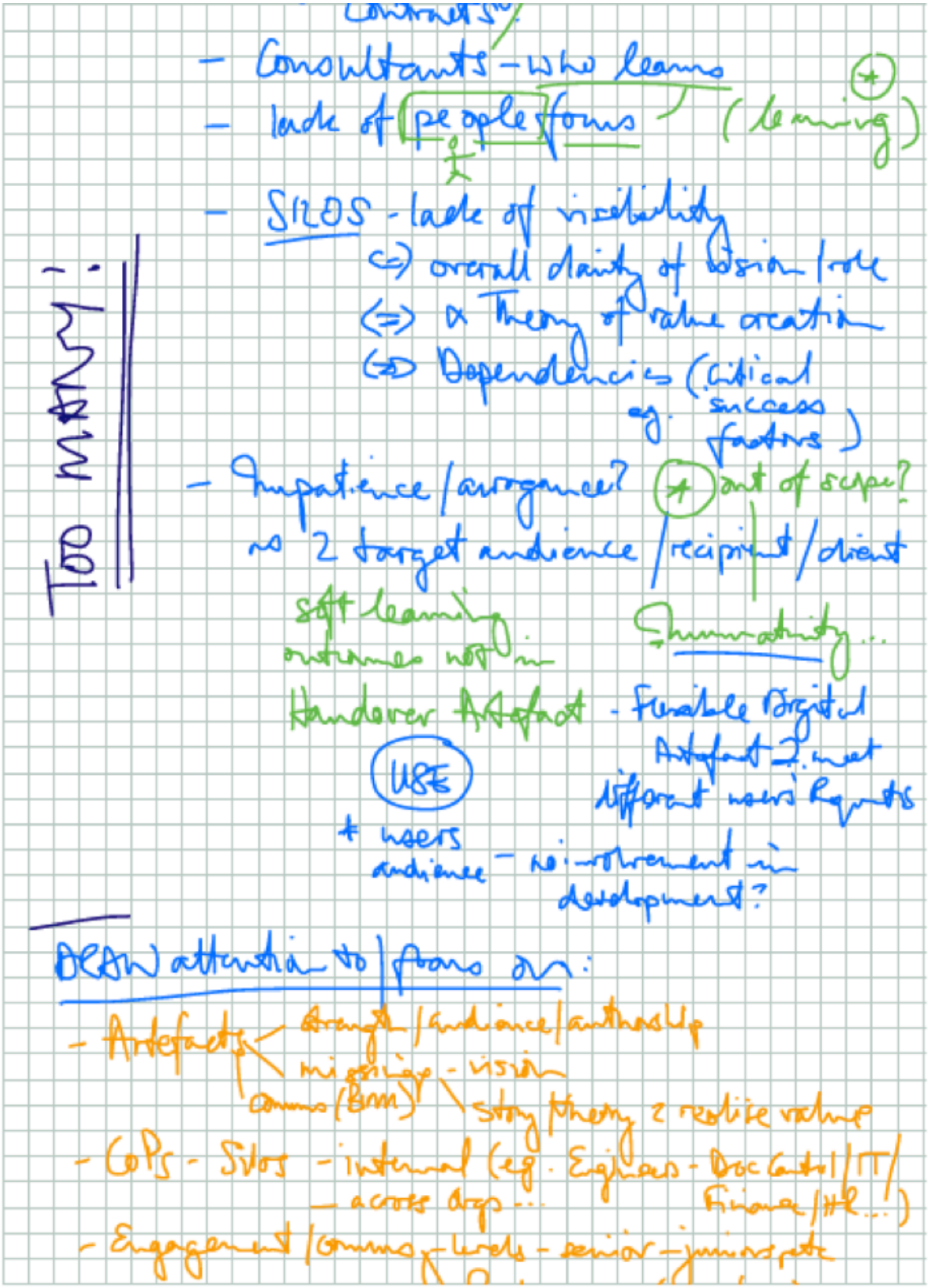
D.7 Monograph planning

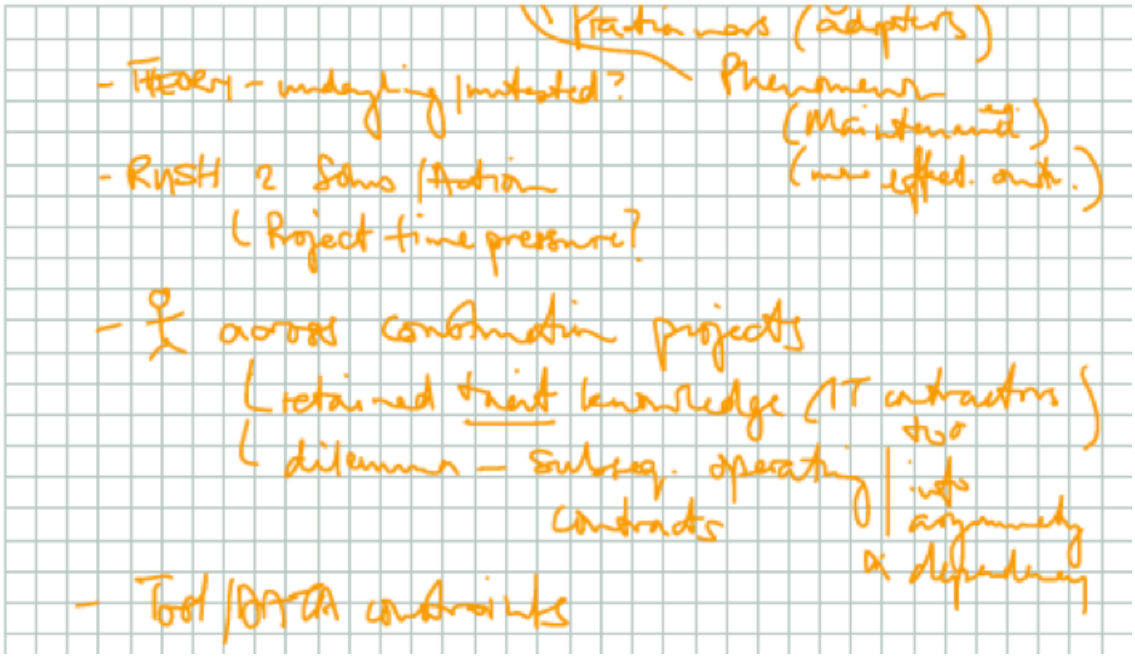
D.7.1 GoCouncil Outline Overview

The following represents a visual brainstorm and potential storyboard worked up for a monogram of the GoCouncil case. It outlines the proposed style, rough structure and inclusions, potential themes and elements to draw reader attention to about the case within a monogram description.



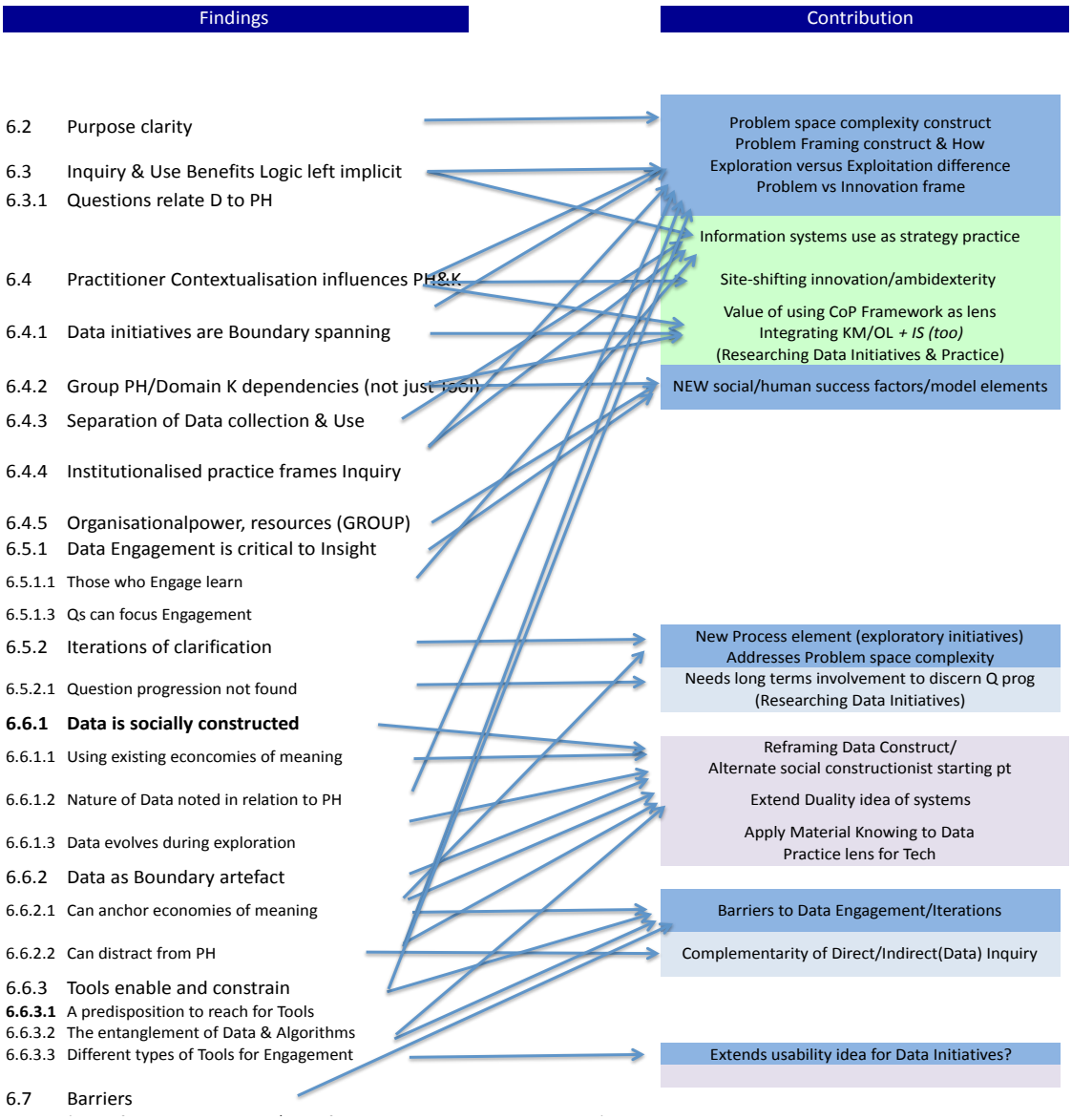
D.7.2 InfraDig Outline Overview

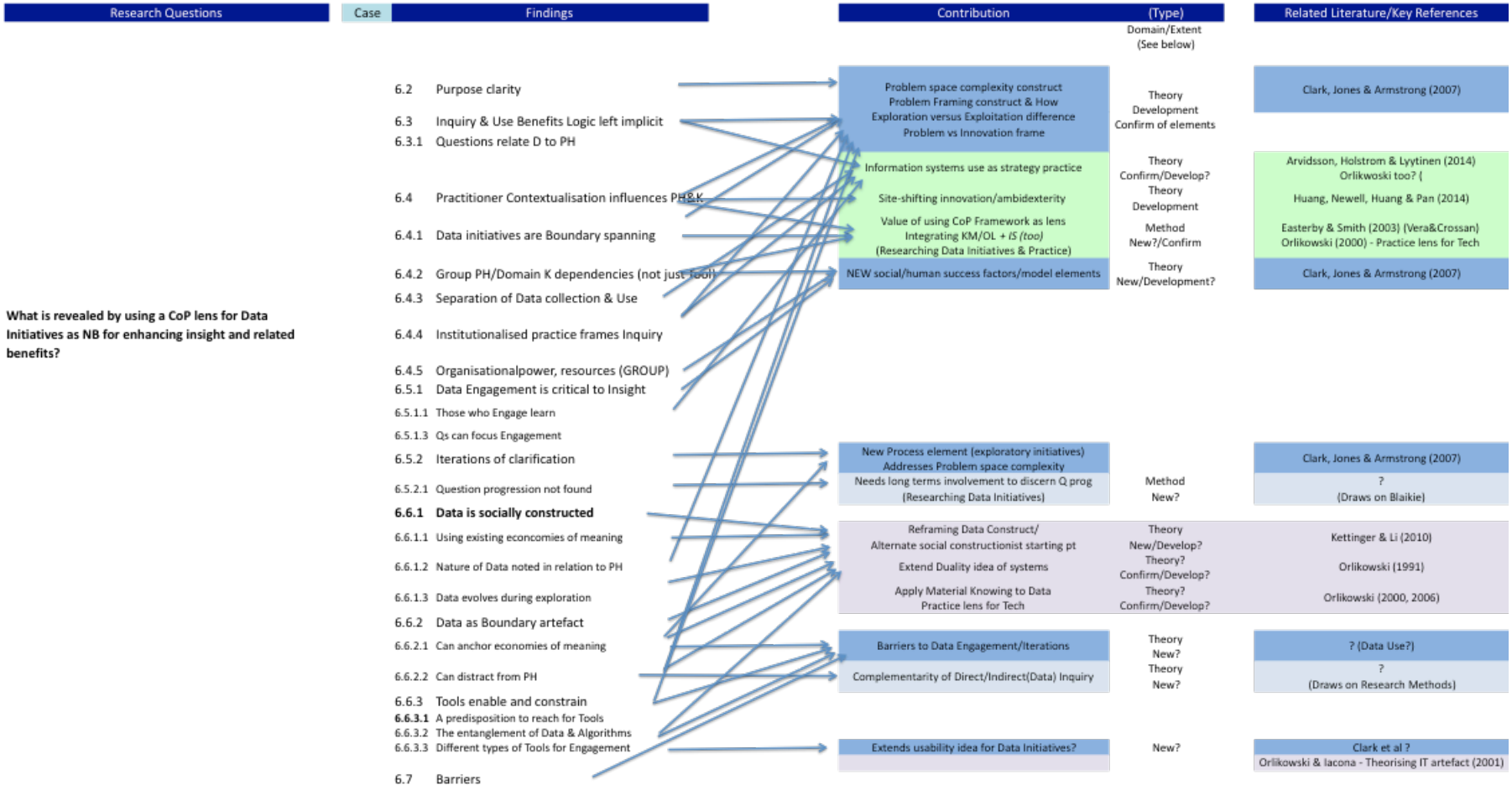

- Consultants - who learns
- lack of people forums (learning)
- SIROS - lack of visibility
⇔ overall clarity of vision/role
⇔ a Theory of value creation
⇔ Dependencies (Critical eg. success factors)
TOO MANY:
- Impatience/arrogance?
no 2 target audience/recipient/client
soft learning outcomes not in Handover Artefact
Immunity...
USE
+ users audience
- no involvement in development?
BEAM attention to/from our:
- Artefacts
- CoPs - SIROS
- Engagement


Practices (adapters)
- THEORY - underlying/mitigated?
Phenomena
(Maintenance)
- RUSH ? Sans Action
(Project time pressure?
- across combination projects
(retained tacit knowledge (IT contractors)
(dilemma - subseq. operating
contracts
two
int
asymmetry
& dependency
- Tool/DATA constraints

D.8 Initial findings mapped to potential themes-contributions

The following outline how a grouped list of findings and observations were initially themed and related to potential theoretical contributions, during initial thesis drafting of findings and ahead of discussing potential contributions during my third review.




Research Questions
Case
Findings
Contribution
(Type)
Domain/Extent
(See below)
Related Literature/Key References
What is revealed by using a CoP lens for Data Initiatives as NB for enhancing insight and related benefits?
6.2 Purpose clarity
6.3 Inquiry & Use Benefits Logic left implicit
6.3.1 Questions relate D to PH
6.4 Practitioner Contextualisation influences PH&K
6.4.1 Data initiatives are Boundary spanning
6.4.2 Group PH/Domain K dependencies (not just Tool)
6.4.3 Separation of Data collection & Use
6.4.4 Institutionalised practice frames Inquiry
6.4.5 Organisationalpower, resources (GROUP)
6.5.1 Data Engagement is critical to Insight
6.5.1.1 Those who Engage learn
6.5.1.3 Qs can focus Engagement
6.5.2 Iterations of clarification
6.5.2.1 Question progression not found
6.6.1 Data is socially constructed
6.6.1.1 Using existing economies of meaning
6.6.1.2 Nature of Data noted in relation to PH
6.6.1.3 Data evolves during exploration
6.6.2 Data as Boundary artefact
6.6.2.1 Can anchor economies of meaning
6.6.2.2 Can distract from PH
6.6.3 Tools enable and constrain
6.6.3.1 A predisposition to reach for Tools
6.6.3.2 The entanglement of Data & Algorithms
6.6.3.3 Different types of Tools for Engagement
6.7 Barriers
Problem space complexity construct
Problem Framing construct & How
Exploration versus Exploitation difference
Problem vs Innovation frame
Theory
Development
Confirm of elements
Clark, Jones & Armstrong (2007)
Information systems use as strategy practice
Theory
Confirm/Develop?
Site-shifting innovation/ambidexterity
Theory
Development
Value of using CoP Framework as lens
Integrating KM/OL + IS (too)
(Researching Data Initiatives & Practice)
Method
New?/Confirm
Arvidsson, Holstrom & Lyytinen (2014)
Orlikwoski too? (
Huang, Newell, Huang & Pan (2014)
Easterby & Smith (2003) (Vera&Crossan)
Orlikowski (2000) - Practice lens for Tech
NEW social/human success factors/model elements
Theory
New/Development?
Clark, Jones & Armstrong (2007)
New Process element (exploratory initiatives)
Addresses Problem space complexity
Clark, Jones & Armstrong (2007)
Needs long terms involvement to discern Q prog
(Researching Data Initiatives)
Method
New?
?
(Draws on Blaikie)
Reframing Data Construct/
Alternate social constructionist starting pt
Theory
New/Develop?
Kettinger & Li (2010)
Extend Duality idea of systems
Theory?
Confirm/Develop?
Orlikowski (1991)
Apply Material Knowing to Data
Practice lens for Tech
Theory?
Confirm/Develop?
Orlikowski (2000, 2006)
Barriers to Data Engagement/Iterations
Theory
New?
? (Data Use?)
Complementarity of Direct/Indirect(Data) Inquiry
Theory
New?
?
(Draws on Research Methods)
Extends usability idea for Data Initiatives?
New?
Clark et al ?
Orlikowski & Iacona - Theorising IT artefact (2001)

Summary Table of contributions

Domains	Extent of Contribution		
	Confirmed	Developed	New findings
Theoretical knowledge	Social aspects of BI projects are important Socially constructed nature of data Various elements of MSS model (e.g. Iterative refinement, tool K, match)	Human reframing and extension of MSS (esp. problem space match/context) Concept of degree of boundary solidity Theorising IT artefacts (disentangling Data & Tools)	**Emerging model for Data Engagement (expl. projects)** **Distinction between human and socio-material** (Data versus Information-Knowledge) (Social) barriers to data engagement
Empirical evidence	No new data type/collection		
Methodology	**Value of adopting CoP lens for data projects (especially 'siting' and bounding human group context)** **Likely value of adopting socio-material lens**	Question-Data 2x2 analysis CoP visual mapping of elements (Prezi) Timeline data plotting/engagement (Prezi)	Limitation of CoP - org structure tension
Knowledge of practice		**Understanding social aspects of data projects (esp. groups involved in data projects & spanning)** **Importance of project 'framing' for outcomes**	

Appendix E – Key CoP and Sensemaking concepts

E.1 Key Communities of Practice Ideas

Reification and Participation

Underpinning his approach is his central tenet about the duality of reification and participation processes and their interaction to produce or negotiate meaning within CoP. His idea of participation is consistent with common usage, i.e. to take part in a common activity or enterprise. He sees this as important in shaping our experiences and in forming communities.

He acknowledges that the term reification is less common and explains it as the process of "giving form to our experience by producing objects that congeal this experience" (p.58), effectively an abstraction of our experience to which we attribute a form of 'excessive' reality, providing a shortcut for communicating. This seems to play a similar role to Tsoukas' (2009) earlier concept of a boundary object (although with facilitating inter-domain dialogue as his focus). Indeed, Wenger acknowledges that reification can refer to both a process and an object, with the latter possibly taking many forms. He also points to a 'double edge' in relation to reification: while it can provide a useful shortcut, often with associated evocative amplification of the related effect and ease of use (as a useful tool), it can become a substitute for a deep understanding of what it represents, especially through its portability and persistence (i.e. they "can take on a life of their own, beyond their context of origin") (p.62).

Having introduced these ideas, he argues strongly for their complementarity, presenting them as an interdependent, fundamental duality, stressing the following characteristics of their interaction – through various illustrations (pp.66-68):

- "Participation and reification are a duality, not opposites"
- "Participation and reification are two dimensions that interact, they do not define a spectrum"

- “Participation and reification imply each other; they do not substitute for each other”
- “Participation and reification transform their relation; they do not translate into each other”
- “Participation and reification describe an interplay; they are not classificatory categories”

Wenger (1998) argues that the benefit of taking this dual process view to arrive at meaning, is that it focuses us on the inherent trade-offs in their complementarity, i.e. “what is reified and what is left to participation” (p.64) or negotiation (effectively). He illustrates this with two extreme examples:

- “A computer program, for instance, could be described as an extreme kind of reification, which can be interpreted by a machine incapable of any participation in its meaning.
- A poem, by contrast, is designed to rely on participation, that is, to maximize the work that the ambiguity inherent in its form can do in the negotiation of meaning.”

He goes on to argue that these dual processes must be in proportion “to compensate for their respective shortcomings” (Wenger: 1998: p.65), highlighting that problems ensue when one is emphasised too much at the expense of the other:

- “If participation prevails – if most of what matters is left unreified – then there may not be enough material to anchor the specificities of coordination and to uncover diverging assumptions. This is why lawyers always want everything in writing.
- If reification prevails – if everything is reified, but with little opportunity for shared experience and interactive negotiation – then there may not be enough overlap in participation to recover a coordinated, relevant, or generative meaning. This helps explain why putting everything in writing does not seem to solve all our problems.” (Wenger: 1998: p.65)

This echoes Tsoukas' (2005) caution about the limits of codified knowledge and may be a useful point of joint departure for characterisation of an IT system or tool as reified knowledge. It may also serve as a possible boundary artefact in order to facilitate dialogue and provide a common context for creating new meaning and knowledge according to Tsoukas (2009).

E.2 Key Sensemaking concepts and ideas (Weick: 1995)

Note: Several original sources cited by Weick have not been reviewed in detail (those not included in the Reference section) but are nevertheless indicated below for transparency and to facilitate seeking them out in his references section

Weick's Sensemaking work contributes several key concepts and ideas in relation to how insights may emerge and my phenomenon of interest, in particular (dealt with in turn in individual sections below):

- The importance of enactment for meaning and the extraction of cues
- The distinction between uncertainty and ambiguity and its implication that more data is only useful when addressing issues of uncertainty rather than ambiguity
- The idea of minimal sensible structures connecting cues with pre-existing frames in order to create meaning
- The interaction of arousal and perceptions of context and its potentially adverse impact on sensemaking

Weick also highlights several concerns about the role of IT in sensemaking, based on the following sensemaking ideas, which are outlined below:

- Enactment
- Ambiguity versus Uncertainty
- Minimal sensible structures

Enactment

His concept of *enactment* is particularly arresting, especially in light of its importance for meaning construction and its recursive nature. Weick starts by explaining his choice to use the word 'enactment' because of its association with law-makers, who he explains "create reality through authoritative acts… they take undefined space, time, and action and draw lines, establish categories, and coin labels that create new features of the environment that did not exist before". (Weick: 1995: p.31). By way of example, he explains that a person's environment is not fixed and independent, instead it is interdependently constructed or framed by their acting in and interacting with their social environment, which then creates constraints and opportunities for them, both constraining and orienting action. He goes on to argue for an extremely important idea: that "self-fulfilling prophecies are the prototype for human sensemaking". "People create and find what they expect to find" (Weick: 1995: p.35). He links this to the idea of people 'oscillating ontologically', having multiple identities and dealing with multiple realities, illustrating this through work by Louis (1980) whom he argues shows that "if … newcomers at first are flooded with surprises, then they start as interpretivists… Over time, as routines develop and the meaning of objects becomes fixed by organizational culture, *facticity* develops as things become taken for granted" (Weick: 1995: p.35). Importantly, he links this process to action and the idea of institutionalisation of both action and meaning through experience, although constantly evolving with points of '*punctuation'*.

Weick's ideas about enactment are very closely interrelated with his sixth characteristic: that it is '*focused on and by extracted cues'*, which may be particularly important for how people create meaning and derive insights. Based on other authors, he points to the role cues play as a reference or starting point to anchor and direct attention, playing a dual role of providing both content and structuring any continuation of attention and meaning creation. He goes on to argue for the overriding importance of context and how it influences the evolution and derivation of meaning from cues in two important ways:

- It impacts on what cues are extracted in the first place (often termed 'search', 'scanning' or 'noticing' by other others. He settles on Goffman (1974)'s 'frame' concept as a "shorthand for the structure of context" (p.51) to encapsulate this idea.

- Secondly, it also impacts on how an extracted cue is interpreted and acted on. Here, to explain, he quotes from social cognition literature that "our attention also orients us to situationally or personally primed categories. Recently, frequently, and chronically encountered categories are more accessible for use, and they profoundly influence the encoding of stimuli" (Fiske & Taylor: 1991: pp.265-266).

He links the latter impact in particular to ethnomethodology's focus on 'indexicals'. He explains that without a context, objects and events can have multiple meanings, while a context collapses these by "providing norms and expectations that constrain explanations" (Weick: 1995: p.53). This seems consistent with Tsoukas' (2005) idea of a narrative structure providing a 'plot' by providing a structure and context. Weick concludes by characterising extracted cues and acting on them as 'acts of faith amid indeterminacy' and that this crucial to facilitating action and sensemaking.

Ambiguity and Uncertainty

Weick draws attention to a central question for data and information use when he clearly distinguishes between ambiguous versus uncertain situations. He starts by looking at ambiguity, drawing on Levine (1985) to characterise it as occurring in instances where meaning is 'unclear' or *equivocal* (i.e. more than one interpretation is possible). He also references Martin (1992), summarising his characterisation as follows: "people judge events to be ambiguous if those events seem to be unclear, highly complex, or paradoxical" (Weick: 1995: p.92). From this starting point, he goes on to argue that the required assumptions to enable rational decision-making are not satisfied in such circumstances. Most importantly, referencing Huber and Daft (1987) and Daft and Lengel (1986) he argues that *more* information (or data) may not resolve such equivocality or confusion and in fact may hinder effective sensemaking, instead arguing for the

central role of language, social interaction and enactment activities to clarify meaning.

By contrast he describes uncertainty as resulting in a 'shock' of *ignorance* rather than *confusion*. Drawing on March (1994) and First, Burns and Stalker (1961) he characterises uncertainty as relating to assessing future consequences relating to current actions, which often occurs when faced with making a choice between different courses of action. Drawing on Frances Milliken (1987)'s definition of uncertainty as "an individual's perceived inability to predict something accurately" (p. 136) he also notes Milliken's observation (in the context of environmental uncertainty) that such uncertainty typically arises from three sources:

- "How components of the environment are changing (state uncertainty), or of
- the impact of environmental changes on the organization (effect uncertainty), or of
- the response options that are open to them (response uncertainty)" (Weick: 1995: p.95)

He goes on to note work by Stinchcombe (1990) that observes that such uncertainty is typically reduced by 'news' and that the nature of residual uncertainty changes over time. Based on work by Daft, Sormunen and Parks (1988) he observes that "occasions for sensemaking should vary as a function of how far into the future a line of action extends, the availability of news, the capability for scanning, the tolerance for risk, the design of the news-collecting structure, and the ease of movement toward sources of news" (Weick: 1995: p.97). He summarises by concluding that "the problem here is not one of too many interpretations, but one of too few… there is an absence of information" (p.97). In such circumstances therefore, more (relevant) information (or data) may indeed be helpful.

The above line of reasoning is implicitly critical of what Weick terms the information-processing perspective, for whom providing more data aids better decision-making. Instead he argues that it is important to first establish if you

are dealing with a situation of uncertainty or ambiguity and to take a different approach to each, i.e. being clear about what kind of sensemaking is required.

This ties into notions of 'problem space complexity' highlighted in the IS literature as important for decision making by Clarke, Jones and Armstrong (2007) in their review of the MSS literature. Weick characterises problems as 'attention-allocation devices' and cites Schön (1991/1983, p.40) as helpful on the importance of problem setting and associated ambiguity in professional work as follows:

> *"although problem setting is a necessary condition for technical problem solving, it is not itself a technical problem. When we set the problem, we select what we will treat as the "things" of the situation, we set the boundaries of our attention to it, and we impose upon it a coherence which allows us to say what is wrong and in what directions the situation needs to be changed. Problem setting is a process in which, interactively we name the things to which we will attend and frame the context in which we will attend to them."*

Minimal sensible structures

For Weick, a crucial component of sensemaking relates to how people "construct roles and interpret objects" (p.109). He argues for "the implicit or explicit operation of some sort of frame (e.g. national culture) within which cues are noticed, extracted, and made sensible" (p.109). He equates this to Starbuck and Milliken (1988)'s idea that "perceptual frameworks categorize data, assign likelihoods to data, hide data, and fill in missing data" (p.51). Weick goes on to argue that "frames and cues can be thought of as vocabularies" (Weick: 1995: p.110) characterising *frames* as more abstract and 'pointing' to *cues*, which he sees as more tangible. He connects this to Upton (1961)'s insight that *for an insight to become meaningful a connection (or a relation) is needed between a frame and a cue*. This seems to echo Polanyi's ideas in relation to tacit knowledge requiring focal attention, peripheral attention and a person to connect them, and such frames are indeed likely often to be tacit rather than

explicit. Weick goes on to argue for the critical importance of such elements and that sensemaking can start with any of the three elements.

The importance of IT in relation to these ideas

Given the pervasiveness of IT within organizations, Weick argues for more research into IT from an interpretive perspective to examine IT's relationship to sensemaking, pointing to work by Orlikowski (1992) and Pentland (1992) as examples of such work. He highlights a growing issue for sensemaking in relation to IT as: "the disparity between the speed and complexity of information technology and the ability of humans to comprehend the outputs of technology. These disparities create the potential for increased arousal" (p.178) with consequent problems that poses for sensemaking. Indeed this can be recognised in the phenomenon of perceived information overload, especially in connection with emails (Hemp: 2009).

Weick goes on to highlight some other problems posed by IT:

- He quotes Lanir (1989) in the context of military systems and distributed decision making, as follows: "Among the problems of such systems are that they underestimate the probabilities of multiple "conditional independent" occurrences" (Weick: 1995: p.178).
- He goes on to highlight that when the 'inconceivable' happens (all too frequently) systems "are unable to reimpose new distinctions on the environment to observe what might be happening when the programmed distinctions break down" (Weick: 1995: p.178).
- Related to the last point, he argues that systems tend to address what is perceived at the design stage (*a priori*) as important and 'controllable' rather than data or information that might be helpful "for improvisation, reframing, or repunctuation" (p.178) which is often simply not available.

He sees these problems as an illustration of the 'strains between decision making and sensemaking', pointing also to the paradox highlighted by Lanir, that while humans can never really act completely rationally, they can

nevertheless design rational systems, which in turn put us under enormous pressure to rely on such systems and perhaps ignore the need for ‘repunctuation’ or reframing when it might be appropriate.

Elsewhere, he characterises the positivist response to this problem as trying to address this through a contingency or boundary conditions approach, through *a priori* anticipation, which he sees as ultimately impossible (i.e. we simply cannot anticipate all eventualities).